

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2004, 14:50:45 ; Search time 56 Seconds  
(without alignments)  
35.318 Million cell updates/sec

Title: US-10-080-100-46

Perfect score: 33  
Sequence: 1 QRANLRA 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: Geneseq1980s:\*  
2: Geneseq1980s:\*  
3: Geneseq2000s:\*  
4: Geneseq2001s:\*  
5: Geneseq2002s:\*  
6: Geneseq2003as:\*  
7: Geneseq2003bs:\*  
8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	100.0	7	5	Abp33146 Zinc fing
2	33	100.0	7	5	Abp33144 Zinc fing
3	33	100.0	7	5	Abu60749 Phage dis
4	33	100.0	7	6	Abu60796 Phage dis
5	33	100.0	8	5	Abp33167 Zinc fing
6	33	100.0	8	6	Abu60713 Phage dis
7	29	87.9	510	2	Aaw70232 Leishmani
8	29	87.9	510	5	Aae24946 Leishmani
9	29	87.9	510	5	Aau71828 Leishmani
10	29	87.9	510	5	Abg60894 Leishmani
11	29	87.9	510	5	Abg71282 L. chagas
12	29	87.9	510	7	Abd78813 Leishmani
13	29	87.9	538	5	Aab71317 L. major
14	29	87.9	538	7	Abd78881 Leishmani
15	29	87.9	1822	5	Abg80407 3-hydroxy
16	28	84.8	15	3	Aab33582 5/15 nati
17	28	84.8	15	3	Aab33581 5/15 nati
18	28	84.8	15	3	Aab23132 Modified
19	28	84.8	15	3	Aab23133 Modified
20	28	84.8	15	4	Aau04795 Modified
21	28	84.8	15	4	Aau04796 Modified
22	28	84.8	124	3	Aag01198 Human sec
23	28	84.8	157	6	Abu52577 Peaut Ar
24	28	84.8	157	6	Abu52575 Peaut Ar
25	28	84.8	166	3	Aab33600 Modified

26	28	84.8	166	4	Aau05035 Modified
27	28	84.8	167	4	Aau04710 Modified
28	28	84.8	269	6	Abm67453 Photorhab
29	28	84.8	440	4	Ag79110 Amino aci
30	28	84.8	440	4	Aab94643 Human pro
31	28	84.8	441	5	Aam52655 Human DNA
32	28	84.8	442	4	Aae11893 Angiogene
33	28	84.8	447	5	Abp62955 Human pol
34	28	84.8	453	4	Ag74830 Human col
35	28	84.8	498	3	Aab08480 Amino aci
36	28	84.8	673	2	Asw81573 Mus dunni
37	28	84.8	1572	5	Abb97562 Novel hum
38	28	84.8	1572	6	Abp98339 Amino aci
39	28	84.8	3038	2	Aar74171 Aspergill
40	27	81.8	77	4	Aao12504 Human pol
41	27	81.8	187	3	Aab38278 Human sec
42	27	81.8	187	3	Aab38276 Human sec
43	27	81.8	229	6	Abm70422 Photorhab
44	27	81.8	231	6	Abu40218 Protein e
45	27	81.8	309	6	Abu39854 Protein e

## ALIGNMENTS

RESULT 1  
ABP33146  
ID ABP33146 standard; peptide; 7 AA.  
XX  
AC ABP33146;  
XX  
DT 12-NOV-2002 (first entry)  
DE  
DE Zinc finger nucleotide binding peptide SEQ ID NO:98.  
XX  
XX Zinc finger binding domain; zinc finger nucleotide binding peptide;  
KW zinc finger; gene expression; modulation; promoter; viral; Lentivirus;  
KW human T-cell lymphocytic virus; HTLV; human immunodeficiency virus; HIV.  
XX  
OS Synthetic.  
XX  
FN WO200266640-A2.  
PN  
PD  
PD 29-AUG-2002.  
XX  
PF 21-FEB-2002; 2002WO-EP001862.  
XX  
PR 21-FEB-2001; 2001US-00791106.

XX (NOVS ) NOVARTIS AG.  
PA (NOVS ) NOVARTIS-ERFINDUNGEN VERW GBS MBH.  
PA (SCRI ) SCRIPPS RES INST.  
XX  
PI Barbas CF, Dreier B;  
XX WPI; 2002-674941/72.

XX New polypeptides comprising 2 to 10 zinc finger-nucleotide binding peptides, useful for modulating gene expression in promoters from viral groups, e.g. Lentivirus group, such as HIV-1 and 2, or HIV-1 and 2.

XX Disclosure; Page 19; 48pp; English.

XX The present invention describes a polypeptide comprising 2 to 12 zinc finger-nucleotide binding peptides in which at least one contains a nucleotide binding region. Also described: (1) an isolated and purified polynucleotide encoding the polypeptide cited above; (2) an expression vector containing the polynucleotide in (1); and (3) a process of regulating expression of a nucleotide sequence that contains the sequence S1 comprising exposing the nucleotide sequence to the polypeptide cited above. (S1) is (5'-ANN)n-3', where n = any integer from 2-12; and N = A, C, G, or T. The polypeptide is useful for modulating gene expression in promoters from viral groups, e.g. Lentivirus group, such as human T-cell

CC Lymphocytic virus (HTLV) 1 and 2, or human immunodeficiency virus (HIV) 1  
 CC or 2. ABQ75687 to ABQ75697 and ABP53112 to ABP53221 represent sequences  
 CC given in the exemplification of the present invention  
 XX  
 XX Sequence 7 AA;

Query Match 100.0%; Score 33; DB 5; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ORANLRA 7  
 Db 1 ORANLRA 7

RESULT 2  
 ABP53144  
 ID ABP53144 standard; peptide; 7 AA.

XX AC ABP53144;  
 XX DT 12-NOV-2002 (first entry)  
 XX DE Zinc finger nucleotide binding peptide SEQ ID NO:46.

XX KW Zinc finger binding domain; zinc finger nucleotide binding peptide;  
 KW zinc finger; gene expression; modulation; promoter; viral; Lentivirus;  
 KW human T-cell lymphocytic virus; HTLV; human immunodeficiency virus; HIV.

XX OS Synthetic.  
 XX PN WO200265640-A2.

XX PD 23-AUG-2002.

XX PF 21-FEB-2002; 2002WO-BP001862.

XX PR 21-FEB-2001; 2001US-00791106.

XX PA (NOVS ) NOVARTIS AG.  
 PA (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.  
 PA (SCRI ) SCRIPPS RES INST.

XX PI Barbas CF, Dreier B;  
 XX DR WPI; 2002-674941/72.

XX PT New polypeptides comprising 2 to 10 zinc finger-nucleotide binding  
 PT peptides, useful for modulating gene expression in promoters from viral  
 PT groups, e.g. Lentivirus group, such HTLV-1 and 2, or HIV-1 and 2.

XX PS Claim 1; Page 19; 48pp; English.

XX CC The present invention describes a polypeptide comprising 2 to 12 zinc  
 CC finger-nucleotide binding peptides in which at least one contains a  
 CC nucleotide binding region. Also described: (1) an isolated and purified  
 CC polynucleotide encoding the polypeptide cited above; (2) an expression  
 CC vector containing the polynucleotide in (1); and (3) a process of  
 CC regulating expression of a nucleotide sequence that contains the sequence  
 CC S1 comprising exposing the nucleotide sequence to the polypeptide cited  
 CC above. (S1) is (5'-ANN)n-3', where n = any integer from 2-12; and N = A,  
 CC C, G, or T. The polypeptide is useful for modulating gene expression in  
 CC promoters from viral groups, e.g. Lentivirus group, such as human T-cell  
 CC lymphocytic virus (HTLV) 1 and 2, or human immunodeficiency virus (HIV) 1  
 CC or 2. ABQ75687 to ABQ75697 and ABP53112 to ABP53221 represent sequences  
 CC given in the exemplification of the present invention

XX SQ Sequence 7 AA;

Query Match 100.0%; Score 33; DB 5; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ORANLRA 7  
 Db 1 ORANLRA 7

RESULT 3  
 ABU60749  
 ID ABU60749 standard; peptide; 7 AA.

XX AC ABU60749;

XX DT 06-MAY-2003 (first entry)

XX DE Phage displayed zinc finger recognising AAA #5.

XX KW Zinc finger; phage display; ant-HIV; virucide; HIV infection;  
 KW human T cell lymphotropic virus infection; HTLV.

XX OS Synthetic.

XX PN US2002165356-A1.

XX PD 07-NOV-2002.

XX PF 21-FEB-2002; 2002US-00080100.

XX PR 21-FEB-2001; 2001US-00367356.

XX PA (SCRI ) SCRIPPS RES INST.

XX PI Barbas CF, Dreier B;

XX DR WPI; 2003-255225/25.

XX PT New zinc finger-nucleotide binding polypeptides, useful in the treatment  
 PT of disorders associated with the aberrant expression or activity of the  
 PT polypeptide, such as HIV infection or human T cell lymphotropic virus  
 PT infection.

XX PS Claim 1; Fig 3; 22pp; English.

XX CC The invention relates a new polypeptide comprising from 2-12 zinc finger-  
 CC nucleotide binding peptides at least one of which contains a nucleotide  
 CC binding region having any of 71 nucleotide binding sequences, fully  
 CC defined in the specification. Also included are an isolated and purified  
 CC polynucleotide that encodes the polypeptide, an expression vector  
 CC containing the polynucleotide and a process of regulating expression of a  
 CC nucleotide sequence that contains the sequence (5'-ANN)n-3', where n is  
 CC an integer from 2-12, the process comprising exposing the nucleotide  
 CC sequence to the zinc finger protein of the invention. The methods and  
 CC compositions of the present invention, are useful for modulating zinc-  
 CC finger protein binding and hence regulating aspects of gene expression.  
 CC The zinc finger polypeptides are useful in the treatment of treatment of  
 CC disorders associated with the aberrant expression or activity of the  
 CC polypeptide, such as viral infection, in particular HIV or human T cell  
 CC lymphotropic virus (HTLV) infection. The present sequence is a synthetic  
 CC nucleotide binding sequence (zinc finger) expressed by a phage display  
 CC library, which may be assembled into the zinc finger protein of the  
 CC invention

XX SQ Sequence 7 AA;

Query Match 100.0%; Score 33; DB 6; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ORANLRA 7  
 Db 1 ORANLRA 7

RESULT 4  
 ABU60796

```

ID ABU60796 standard; peptide; 7 AA.
XX
AC ABU60796;
XX
DT 06-MAY-2003 (first entry)
XX
DE Phage displayed zinc finger recognising AAA #6.
XX
KW Zinc finger; phage display; ant-HIV; virucide; HIV infection;
KW human T cell lymphotropic virus infection; HTLV.
XX
OS Synthetic.
XX
PN US2002165356-A1.
XX
PD 07-NOV-2002.
XX
PF 21-FEB-2002; 2002US-00080100.
XX
PR 21-FEB-2001; 2001US-00367356.
XX
PA (SCRI ) SCRIPPS RES INST.
XX
PI Barbas CF, Dreier B;
XX
DR WPI; 2003-255225/25.
XX
PT New zinc finger-nucleotide binding polypeptides, useful in the treatment
PT of disorders associated with the aberrant expression or activity of the
PT polypeptide, such as HIV infection or human T cell lymphotropic virus
PT infection.
XX
PS Disclosure; Page 7; 22pp; English.
XX
CC The invention relates a new polypeptide comprising from 2-12 zinc finger-
CC nucleotide binding peptides at least one of which contains a nucleotide
CC binding region having any of 71 nucleotide binding sequences, fully
CC defined in the specification. Also included are an isolated and purified
CC polynucleotide that encodes the polypeptide, an expression vector
CC containing the polynucleotide and a process of regulating expression of a
CC nucleotide sequence that contains the sequence (5'-ANN)n-3', where n is
CC an integer from 2-12, the process comprising exposing the nucleotide
CC sequence to the zinc finger protein of the invention. The methods and
CC compositions of the present invention, are useful for modulating zinc-
CC finger protein binding and hence regulating aspects of gene expression.
CC The zinc finger polypeptides are useful in the treatment of treatment of
CC disorders associated with the aberrant expression or activity of the
CC polypeptide, such as viral infection, in particular HIV or human T cell
CC lymphotropic virus (HTLV) infection. The present sequence is a synthetic
CC nucleotide binding sequence (zinc finger) expressed by a phage display
CC library, which may be assembled into the zinc finger protein of the
CC invention
XX
SQ Sequence 7 AA;
Query Match 100.0%; Score 33; DB 6; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QRANLRA 7
DB 1 QRANLRA 7
RESULT 5
ABP53167
ID ABP53167 standard; peptide; 8 AA.
XX
AC ABP53167;
XX
DT 12-NOV-2002 (first entry)
XX
DE Zinc finger nucleotide binding peptide SEQ ID NO:10.

```

```

XX Zinc finger binding domain; zinc finger nucleotide binding peptide;
KW zinc finger; gene expression; modulation; promoter; viral; Lentivirus;
KW human T-cell lymphocytic virus; HTLV; human immunodeficiency virus; HIV.
XX
OS Synthetic.
XX
PN WO200266640-A2.
XX
PD 29-AUG-2002.
XX
PF 21-FEB-2002; 2002WO-EP001862.
XX
PR 21-FEB-2001; 2001US-00791106.
XX
PA (NOVS ) NOVARTIS AG.
PA (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.
PA (SCRI ) SCRIPPS RES INST.
XX
PI Barbas CF, Dreier B;
XX
DR WPI; 2002-674941/72.
XX
PT New polypeptides comprising 2 to 10 zinc finger-nucleotide binding
PT peptides, useful for modulating gene expression in promoters from viral
PT groups, e.g. Lentivirus group, such HTLV-1 and 2, or HIV-1 and 2.
XX
PS Claim 1; Fig 2; 48pp; English.
XX
CC The present invention describes a polypeptide comprising 2 to 12 zinc
CC finger-nucleotide binding peptides in which at least one contains a
CC nucleotide binding region. Also described: (1) an isolated and purified
CC polynucleotide encoding the polypeptide cited above; (2) an expression
CC vector containing the polynucleotide in (1); and (3) a process of
CC regulating expression of a nucleotide sequence that contains the sequence
CC S1 comprising exposing the nucleotide sequence to the polypeptide cited
CC above. (S1) is (5'-ANN)n-3', where n = any integer from 2-12; and N = A,
CC C, G, or T. The polypeptide is useful for modulating gene expression in
CC promoters from viral groups, e.g. Lentivirus group, such as human T-cell
CC lymphocytic virus (HTLV) 1 and 2, or human immunodeficiency virus (HIV) 1
CC or 2. AB075697 to AB075697 and ABP53112 to ABP53221 represent sequences
CC given in the exemplification of the present invention
XX
SQ Sequence 8 AA;
Query Match 100.0%; Score 33; DB 5; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QRANLRA 7
DB 2 QRANLRA 8
RESULT 6
ABU60713
ID ABU60713 standard; peptide; 8 AA.
XX
AC ABU60713;
XX
DT 06-MAY-2003 (first entry)
XX
DE Phage displayed zinc finger recognising AAA #4.
XX
KW Zinc finger; phage display; ant-HIV; virucide; HIV infection;
KW human T cell lymphotropic virus infection; HTLV.
XX
OS Synthetic.
XX
PN US2002165356-A1.
XX
PD 07-NOV-2002.
XX

```

PF 21-FEB-2002; 2002US-00080100.  
 XX 21-FEB-2001; 2001US-00367356.  
 PR (SCRI ) SCRIPPS RES INST.  
 XX Barbas CF, Dreier B;  
 PI WPI; 2003-255225/25.  
 XX New zinc finger-nucleotide binding polypeptides, useful in the treatment  
 PT of disorders associated with the aberrant expression or activity of the  
 PT polypeptide, such as HIV infection or human T cell lymphotropic virus  
 PT infection.  
 XX Claim 1; Fig 2; 22pp; English.  
 PS The invention relates a new polypeptide comprising from 2-12 zinc finger-  
 CC nucleotide binding peptides at least one of which contains a nucleotide  
 CC binding region having any of 71 nucleotide binding sequences, fully  
 CC defined in the specification. Also included are an isolated and purified  
 CC polynucleotide that encodes the polypeptide, an expression vector  
 CC containing the polynucleotide and a process of regulating expression of a  
 CC nucleotide sequence that contains the sequence (5'-ANN)n-3', where n is  
 CC an integer from 2-12, the process comprising exposing the nucleotide  
 CC sequence to the zinc finger protein of the invention. The methods and  
 CC compositions of the present invention, are useful for modulating zinc-  
 CC finger protein binding and hence regulating aspects of gene expression.  
 CC The zinc finger polypeptides are useful in the treatment of treatment of  
 CC disorders associated with the aberrant expression or activity of the  
 CC polypeptide, such as viral infection, in particular HIV or human T cell  
 CC lymphotropic virus (HTLV) infection. The present sequence is a synthetic  
 CC nucleotide binding sequence (zinc finger) expressed by a phage display  
 CC library, which may be assembled into the zinc finger protein of the  
 CC invention  
 XX  
 SQ Sequence 8 AA;  
 Query Match 100.0%; Score 33; DB 6; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QRANLRA 7  
 DB |||||  
 2 QRANLRA 8  
 RESULT 7  
 ID AAW70232  
 XX AAW70232 standard; protein; 510 AA.  
 AC AAW70232;  
 DT 17-OCT-2003 (revised)  
 DT 13-NOV-1998 (first entry)  
 XX Leishmania antigen LcgSP3 protein.  
 DE Leishmania antigen; immune response; infection detection; therapy;  
 KW humoral response induction; cellular response induction; cancer;  
 KW interleukin-12 production.  
 XX Leishmania donovani chagasi.  
 OS Leishmania donovani chagasi.  
 XX WO9835045-A2.  
 PN 13-AUG-1998.  
 PD 12-FEB-1998; 98WO-US003002.  
 XX 12-FEB-1997; 97US-00798841.  
 PF 12-FEB-1997; 97US-00798841.  
 PR 27-AUG-1997; 97US-00920609.  
 PT New polypeptide comprising immunogenic part of Leishmania antigen, useful

PA (CORI-) CORIXA CORP.  
 XX Reed SG, Campos-Neto A, Webb JR, Dillon DC, Skeiky YA;  
 XX WPI; 1998-447242/38.  
 DR N-PSDB; AAV47577.  
 XX New immunogenic fragments of Leishmania antigens and related nucleic  
 PT acid, vectors and host cells - are useful for diagnosis, prevention and  
 PT treatment of leishmaniasis, also to induce production of interleukin-12  
 PT generally.  
 XX Claim 13; Page 125-127; 194pp; English.  
 PS This sequence encodes a Leishmania antigen (LAG) of the invention,  
 CC designated LcgSP3. Compositions and vaccines containing the protein are  
 CC used to generate a protective or therapeutic immune response against the  
 CC Leishmania species donovani, chagasi, infantum, major, amazonensis,  
 CC braziliensis, panamensis, tropica or guayanaensis. They can also be used  
 CC to detect infection (in a skin test). The compositions induce a humoral  
 CC and/or cellular response, specifically of Th1 type, particularly  
 CC including induction of interleukin-12 (IL-12) production. They may thus  
 CC be used more generally to treat any condition (e.g. bacterial, viral or  
 CC protozoal infection, or cancer) which responds to IL-12. (Updated on 17-  
 CC OCT-2003 to standardise OS field)  
 XX Sequence 510 AA;  
 SQ Query Match 87.9%; Score 29; DB 2; Length 510;  
 Best Local Similarity 100.0%; Pred. No. 3.6e-02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QRANLRA 6  
 DB |||||  
 114 QRANLRA 119  
 RESULT 8  
 ID AAE24946 standard; protein; 510 AA.  
 XX AAE24946;  
 AC AAE24946;  
 XX 29-AUG-2003 (revised)  
 DT 22-OCT-2002 (first entry)  
 XX Leishmania chagasi LcgSP3 antigenic protein.  
 DE Immunogen; Leishmania antigen; therapy; delayed-type hypersensitivity;  
 KW leishmaniasis; vaccine; interleukin-12 stimulation; cancer; protozoacide;  
 KW virucide; bactericide; cytostatic; immune response; LcgSP3 protein.  
 XX Leishmania donovani chagasi.  
 OS Leishmania donovani chagasi.  
 XX US6375955-B1.  
 PN 23-APR-2002.  
 PD 12-FEB-1998; 98US-00022765.  
 PF 22-SEP-1995; 95US-00533669.  
 PR 12-FEB-1997; 97US-00798841.  
 PR 27-AUG-1997; 97US-00920609.  
 XX (CORI-) CORIXA CORP.  
 PA Reed SG, Campos-Neto A, Webb JR, Dillon DC, Skeiky YAW, Bhatia A;  
 PI Probst P;  
 XX WPI; 2002-424754/45.  
 DR N-PSDB; AAD40304.  
 XX New polypeptide comprising immunogenic part of Leishmania antigen, useful

PT for treatment, prevention and diagnosis of leishmaniasis.  
 XX Example 14; Col 121-124; 111pp; English.  
 XX The present invention relates to novel proteins comprising immunogenic part of Leishmania antigen. Compositions containing sequences of the invention and other Leishmania antigen related polypeptides are useful for preventing, treating and detecting (in delayed-type hypersensitivity skin tests) leishmaniasis. They can also be used to treat any diseases CC responsive to interleukin-12 stimulation, including bacterial, viral and CC protozoal infections and cancer. Sequences of the invention are useful as CC vaccines. The present sequence is Leishmania chagasi LcgSP3 antigenic CC protein. (Updated on 29-AUG-2003 to standardise OS field)  
 XX Sequence 510 AA;  
 SQ

Query Match 87.9%; Score 29; DB 5; Length 510;  
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QRANLR 6  
 DB 114 QRANLR 119

RESULT 9  
 AAU71828  
 ID AAU71828 standard; protein; 510 AA.  
 XX AC AAU71828;  
 XX DT 29-AUG-2003 (revised)  
 XX DT 26-FEB-2002 (first entry)  
 XX DE Leishmania antigen LcgSP3.  
 XX Leishmaniasis; Leishmania antigen; immunostimulant; protozoacide; ldp23;  
 KW interleukin-15; LbhsP83; M15; Lt-1; LbeIF4A; LmSP1a; LmSP9a; MAPS-1A;  
 KW LmgSP1; LmgSP3; LmgSP8; LmgSP9; LmgSP13; LmgSP19; LcgSP1; LcgSP3;  
 KW LcgSP4; LcgSP8; LcgSP10; 1G6-34; 1E6-44; 4A5-63; 1B11-39; 2A10-37;  
 KW 4G2-83; 4H6-41; 8G3-100.  
 XX OS Leishmania donovani chagasi.  
 XX WO200179276-A2.  
 XX PD 25-OCT-2001.  
 XX PF 05-APR-2001; 2001WO-US011254.  
 XX PR 14-APR-2000; 2000US-00551974.  
 XX PR 05-MAY-2000; 2000US-00565501.  
 XX PR 14-AUG-2000; 2000US-00639206.  
 XX PA (CORI-) CORIXA CORP.  
 XX Reed SG, Campos-Neto A, Webb JR, Dillon DC, Skeiky YAW, Bhatia A;  
 PI Coler RM, Probst P;  
 XX WPI: 2002-061971/08.  
 XX DR N-PSDB; AAS96041.  
 XX New isolated Leishmania antigens, useful for prevention, treatment and PT diagnosis of leishmaniasis, also related nucleic acids for genetic PT vaccination.  
 XX Claim 1; Page 143-144; 193pp; English.  
 XX The invention relates to polypeptides comprising an immunogenic part of a CC Leishmania antigen. The Leishmania polypeptides and their associated DNA CC sequences, epitopes and fusion proteins are used in the production of CC compositions used for inducing a protective immune response against CC leishmaniasis, for prevention and treatment of the disease. The

CC compositions can also be used generally to treat diseases that respond to interleukin-15 stimulation. In addition, the products may contain an immunostimulant. Sequences AAU71805-AAU71862 represent Leishmania antigens and antigenic peptides of the invention. (Updated on 29-AUG-2003 to standardise OS field)  
 XX Sequence 510 AA;  
 SQ

Query Match 87.9%; Score 29; DB 5; Length 510;  
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QRANLR 6  
 DB 114 QRANLR 119

RESULT 10  
 ABG60884  
 ID ABG60884 standard; protein; 510 AA.  
 XX AC ABG60884;  
 XX DT 29-AUG-2003 (revised)  
 XX DT 19-AUG-2002 (first entry)  
 XX DE Leishmania antigenic polypeptide #21.  
 XX Leishmania antigen; Th1; interleukin-12; IL-12; Leishmaniasis;  
 KW protozoacide; immune response enhancer; vaccine.  
 XX OS Leishmania donovani chagasi.  
 XX US6365165-B1.  
 XX PD 02-APR-2002.  
 XX PF 30-OCT-1998; 98US-00183861.  
 XX PR 22-SEP-1995; 95US-00533669.  
 XX PR 12-FEB-1997; 97US-00798841.  
 XX PR 27-AUG-1997; 97US-00920809.  
 XX PR 12-FEB-1998; 98US-00022765.  
 XX PA (CORI-) CORIXA CORP.  
 XX Reed SG, Campos-Neto A, Webb JR, Dillon DC, Skeiky YAW;  
 XX WPI: 2002-433419/46.  
 XX DR N-PSDB; ABK81750.  
 XX Stimulating immune responses in a patient for preventing or treating PT Leishmaniasis, by administering a vaccine comprising a polypeptide PT comprising an immunogenic portion or epitope of Leishmania antigen.  
 XX Claim 1; Col 123-126; 114pp; English.  
 XX The invention relates to stimulating an immune response in a patient CC comprising administering to the patient a vaccine containing a non- CC specific immune response enhancer and a polypeptide having an immunogenic CC portion of Leishmania antigen, or a polypeptide having two contiguous CC epitopes of a Leishmania antigen. The method is useful for stimulating an CC immune response, in particular a Th1 response or interleukin-12 (IL-12) CC production in a patient, for preventing and treating Leishmaniasis. CC Sequences ABG60864-ABG60896 and ABG60903-ABG60910 represent Leishmania CC antigenic polypeptides and peptide fragments of the invention. (Updated CC on 29-AUG-2003 to standardise OS field)  
 XX Sequence 510 AA;  
 SQ

Query Match 87.9%; Score 29; DB 5; Length 510;  
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRANLR 6  
 DB 114 QRANLR 119

Query Match 87.9%; Score 29; DB 5; Length 510;  
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRANLR 6  
 DB 114 QRANLR 119

RESULT 12  
 ADB78813  
 ID ADB78813 standard; protein; 510 AA.  
 XX  
 AC ADE78813;  
 XX  
 DT 04-DEC-2003 (first entry)  
 DE Leishmania antigen LcgSP3.  
 XX  
 KW Antigen; protozoacide; antibacterial; virucide; cytostatic;  
 KW immunostimulant; leishmaniasis; Leishmania infection; immune response;  
 KW interleukin-2 stimulation; cancer; bacterial infection; viral infection;  
 KW protozoan infection.  
 XX  
 OS Leishmania donovani chagasi.  
 XX  
 PN US2002169285-A1.  
 XX  
 PD 14-NOV-2002.  
 XX  
 PF 20-NOV-2001; 2001US-00991496.  
 XX  
 PR 22-SEP-1995; 95US-00533669.  
 PR 12-FEB-1997; 97US-00798841.  
 PR 27-AUG-1997; 97US-00920609.  
 PR 12-FEB-1998; 98US-00022765.  
 PR 30-OCT-1998; 98US-00183861.  
 PR 14-APR-2000; 2000US-00551974.  
 PR 05-MAY-2000; 2000US-00565501.  
 PR 14-AUG-2000; 2000US-00639206.  
 PR 04-JUN-2001; 2001US-00874923.  
 XX  
 PA (REED/) REED S G.  
 PA (CAMP/) CAMPOS-NETO A.  
 PA (WEBB/) WEBB J R.  
 PA (DILL/) DILLON D C.  
 PA (BHAT/) BHATIA A.  
 PA (COLE/) COLER R N.  
 PA (PROB/) PROBST P.  
 PA (BRAN/) BRANNON M.  
 XX  
 PI Reed SG, Campos-Neto A, Webb JR, Dillon DC, Skeiky YAW, Bhatia A;  
 PI Coler RN, Probst P, Brannon M;  
 XX  
 DR WPI; 2002-635457/68.  
 DR N-PSDB; AAR88543.  
 XX  
 PT New polypeptide containing at least an immunogenic portion of one or more  
 PT Leishmania antigens or their variants, useful for preventing, treating  
 PT and detecting leishmaniasis, and stimulating immune responses in  
 PT patients.  
 XX  
 PS Claim 1; Page 60-61; 163pp; English.  
 CC This invention describes a novel polypeptide containing an immunogenic  
 CC portion of a Leishmania antigen or its variant which has antiparasitic  
 CC and immunostimulant activity. The compositions and methods of the present  
 CC invention are useful for preventing, treating and detecting  
 CC leishmaniasis, and stimulating immune responses in patients against  
 CC leishmaniasis. The polypeptides and the polynucleotides encoding them can  
 CC be used for gene therapy, in vaccines or as interleukin-12 agonists. The  
 CC compositions and methods of the present invention as compared to prior  
 CC art, are more improved therapeutic modalities in the diagnosis,  
 CC prevention and treatment of leishmaniasis. (Updated on 29-AUG-2003 to  
 CC standardise OS field)  
 XX  
 SQ Sequence 510 AA;

QY 1 QRANLR 6  
 DB 114 QRANLR 119

Query Match 87.9%; Score 29; DB 5; Length 510;  
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRANLR 6  
 DB 114 QRANLR 119

RESULT 12  
 ADB78813  
 ID ADB78813 standard; protein; 510 AA.  
 XX  
 AC ADE78813;  
 XX  
 DT 04-DEC-2003 (first entry)  
 DE Leishmania antigen LcgSP3.  
 XX  
 KW Antigen; protozoacide; antibacterial; virucide; cytostatic;  
 KW immunostimulant; leishmaniasis; Leishmania infection; immune response;  
 KW interleukin-2 stimulation; cancer; bacterial infection; viral infection;  
 KW protozoan infection.  
 XX  
 OS Leishmania donovani chagasi.  
 XX  
 PN US2002169285-A1.  
 XX  
 PD 14-NOV-2002.  
 XX  
 PF 20-NOV-2001; 2001US-00991496.  
 XX  
 PR 22-SEP-1995; 95US-00533669.  
 PR 12-FEB-1997; 97US-00798841.  
 PR 27-AUG-1997; 97US-00920609.  
 PR 12-FEB-1998; 98US-00022765.  
 PR 30-OCT-1998; 98US-00183861.  
 PR 14-APR-2000; 2000US-00551974.  
 PR 05-MAY-2000; 2000US-00565501.  
 PR 14-AUG-2000; 2000US-00639206.  
 PR 04-JUN-2001; 2001US-00874923.  
 XX  
 PA (REED/) REED S G.  
 PA (CAMP/) CAMPOS-NETO A.  
 PA (WEBB/) WEBB J R.  
 PA (DILL/) DILLON D C.  
 PA (BHAT/) BHATIA A.  
 PA (COLE/) COLER R N.  
 PA (PROB/) PROBST P.  
 PA (BRAN/) BRANNON M.  
 XX  
 PI Reed SG, Campos-Neto A, Webb JR, Dillon DC, Skeiky YAW, Bhatia A;  
 PI Coler RN, Probst P, Brannon M;  
 XX  
 DR WPI; 2002-635457/68.  
 DR N-PSDB; AAR88543.  
 XX  
 PT New polypeptide containing at least an immunogenic portion of one or more  
 PT Leishmania antigens or their variants, useful for preventing, treating  
 PT and detecting leishmaniasis, and stimulating immune responses in  
 PT patients.  
 XX  
 PS Claim 1; Page 60-61; 163pp; English.  
 CC This invention describes a novel polypeptide containing an immunogenic  
 CC portion of a Leishmania antigen or its variant which has antiparasitic  
 CC and immunostimulant activity. The compositions and methods of the present  
 CC invention are useful for preventing, treating and detecting  
 CC leishmaniasis, and stimulating immune responses in patients against  
 CC leishmaniasis. The polypeptides and the polynucleotides encoding them can  
 CC be used for gene therapy, in vaccines or as interleukin-12 agonists. The  
 CC compositions and methods of the present invention as compared to prior  
 CC art, are more improved therapeutic modalities in the diagnosis,  
 CC prevention and treatment of leishmaniasis. (Updated on 29-AUG-2003 to  
 CC standardise OS field)  
 XX  
 SQ Sequence 510 AA;

CC patient, or for treating a patient with a disease responsive to  
CC interleukin (IL)-2 stimulation, where the disease is cancer or an  
CC infection such as bacterial, viral or protozoan infection. The antigen is  
CC useful for preventing or treating leishmaniasis. The present sequence  
CC represents a leishmania antigen (or fragment).

XX  
SQ Sequence 510 AA;  
Query Match 87.9%; Score 29; DB 7; Length 510;  
Best Local Similarity 100.0%; Pred. No. 3.6e+02; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRANLR 6  
Db 114 QRANLR 119  
|||  
|||

RESULT 13  
AAB71317  
ID AAB71317 standard; protein; 538 AA.  
XX  
AC AAB71317;  
XX  
DT 29-AUG-2003 (revised)  
DT 18-NOV-2002 (first entry)  
XX  
DE L. major and L. chagasi associated antigen SEQ ID 118.  
XX  
XX Antigen; immunogenic; antiparasitic; immunostimulant; leishmaniasis;  
KW gene therapy; vaccine; interleukin-12 agonist.  
XX  
OS Leishmania major.  
OS Leishmania donovani chagasi.  
XX  
PN US2002081320-A1.  
XX  
XX 27-JUN-2002.  
XX  
PF 04-JUN-2001; 2001US-00874923.  
XX  
PR 22-SEP-1995; 95US-00533669.  
PR 12-FEB-1997; 97US-00798841.  
PR 27-AUG-1997; 97US-00920609.  
PR 12-FEB-1998; 98US-00022765.  
PR 30-OCT-1998; 98US-00183861.  
PR 14-APR-2000; 2000US-00551974.  
PR 05-MAY-2000; 2000US-00565501.  
PR 14-AUG-2000; 2000US-00639206.  
XX  
PA (REED/) REED S G.  
PA (CAMP/) CAMPOS-NETO A.  
PA (WEBB/) WEBB J R.  
PA (DILL/) DILLON D C.  
PA (SKEI/) SKEIKY Y A W.  
PA (BHAT/) BHATIA A.  
PA (COLE/) COLER R N.  
PA (PROB/) PROBST P.  
PA (BRAN/) BRANNON M.  
XX  
PI Reed SG, Campos-Neto A, Webb JR, Dillon DC, Skeiky YAW, Bhatia A;  
PI Coler RN, Probst P, Brannon M;  
XX  
DR WPI; 2002-635457/68.  
DR N-PSDB; AAF88575.  
XX  
XX New polypeptide containing at least an immunogenic portion of one or more  
PT Leishmania antigens or their variants, useful for preventing, treating  
PT and detecting leishmaniasis, and stimulating immune responses in  
PT patients.  
XX  
XX Claim 1; Page 118-119; 163pp; English.  
PS This invention describes a novel polypeptide containing an immunogenic  
CC

CC portion of a Leishmania antigen or its variant which has antiparasitic  
CC and immunostimulant activity. The compositions and methods of the present  
CC invention are useful for preventing, treating and detecting  
CC leishmaniasis, and stimulating immune responses in patients against  
CC leishmaniasis. The polypeptides and the polynucleotides encoding them can  
CC be used for gene therapy, in vaccines or as interleukin-12 agonists. The  
CC compositions and methods of the present invention, as compared to prior  
CC art, are more improved therapeutic modalities in the diagnosis,  
CC prevention and treatment of leishmaniasis. (Updated on 29-AUG-2003 to  
CC standardise OS field)

XX  
SQ Sequence 538 AA;  
Query Match 87.9%; Score 29; DB 5; Length 538;  
Best Local Similarity 100.0%; Pred. No. 3.8e+02; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRANLR 6  
Db 117 QRANLR 122  
|||  
|||

RESULT 14  
ADB78881  
ID ADB78881 standard; protein; 538 AA.  
XX  
AC ADB78881;  
XX  
DT 04-DEC-2003 (first entry)  
DE Leishmania antigen LcgSP3 full length protein.  
XX  
XX Antigen; protozoacide; antibacterial; virucide; cytostatic;  
KW immunostimulant; leishmaniasis; Leishmania infection; immune response;  
KW interleukin-2 stimulation; cancer; bacterial infection; viral infection;  
KW protozoan infection.  
XX  
OS Leishmania major.  
OS Leishmania donovani chagasi.  
XX  
PN US2002169285-A1.  
XX  
PD 14-NOV-2002.  
XX  
PF 20-NOV-2001; 2001US-00991496.  
XX  
PR 22-SEP-1995; 95US-00533669.  
PR 12-FEB-1997; 97US-00798841.  
PR 27-AUG-1997; 97US-00920609.  
PR 12-FEB-1998; 98US-00022765.  
PR 30-OCT-1998; 98US-00183861.  
PR 14-APR-2000; 2000US-00551974.  
PR 05-MAY-2000; 2000US-00565501.  
PR 14-AUG-2000; 2000US-00639206.  
PR 04-JUN-2001; 2001US-00874923.  
XX  
PA (REED/) REED S G.  
PA (CAMP/) CAMPOS-NETO A.  
PA (WEBB/) WEBB J R.  
PA (DILL/) DILLON D C.  
XX  
PI Reed SG, Campos-Neto A, Webb JR, Dillon DC;  
XX  
DR WPI; 2003-605673/57.  
DR N-PSDB; ADB78876.  
XX  
XX Novel isolated polypeptide useful for preventing or treating  
PT leishmaniasis, comprises an immunogenic portion of a Leishmania antigen  
PT or its variant.  
XX  
XX Claim 1; Page 120-121; 183pp; English.  
PS The invention relates to an isolated polypeptide comprising an  
CC

CC immunogenic portion of a Leishmania antigen or its. Also included are  
 CC antigenic epitopes, fusion proteins comprising an isolated polypeptide  
 CC fusion protein comprising at least two contiguous antigenic epitopes,  
 CC polynucleotides encoding the antigens or fusion proteins, a recombinant  
 CC expression vector comprising the polynucleotide, a host cell transformed  
 CC with the vector and a composition (pharmaceutical or immunogenic)  
 CC comprising the antigen or fusion protein and a physiologically acceptable  
 CC carrier. The compositions are useful for inducing protective immunity  
 CC against leishmaniasis in a patient. The fusion protein is useful for  
 CC detecting Leishmania infection in a patient by contacting dermal cells of  
 CC the patient with the composition and detecting an immune response on the  
 CC patient's skin, where the immune response is induration. The compositions  
 CC are useful for stimulating a cellular and/or humoral immune response in a  
 CC patient, or for treating a patient with a disease responsive to  
 CC interleukin (IL)-2 stimulation, where the disease is cancer or an  
 CC infection such as bacterial, viral or protozoan infection. The antigen is  
 CC useful for preventing or treating leishmaniasis. The present sequence  
 CC represents a Leishmania antigen (or fragment).

XX SQ Sequence 538 AA;

Query Match 87.9%; Score 29; DB 7; Length 538;

Best Local Similarity 100.0%; Pred. No. 3.8e-02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRANLR 6

DB 117 QRANLR 122

RESULT 15

ABG80407

ID ABG80407 standard; protein; 1822 AA.

AC ABG80407;

DT 29-NOV-2002 (first entry)

XX 3-hydroxypropionic acid production related polypeptide #3.

DE 3-Hydroxypropionic acid; 3-HP; polymerised 3-HP; 3-HP ester;

KW polymerised acrylate; acrylate ester; lactyl CoA-dehydratase;

KW 3-hydroxypropionyl-CoA dehydratase; food; feed; preservative.

XX Unidentified.

OS WO200242418-A2.

XX 10-MAY-2002.

PN 20-NOV-2001; 2001WO-US043607.

XX 20-NOV-2000; 2000US-0252123P.

XX 20-APR-2001; 2001US-0285478P.

PR 20-JUL-2001; 2001US-0306727P.

PR 07-SEP-2001; 2001US-0317845P.

XX (CRGI ) CARGILL INC.

PA Gokarn RR, Selifonova OV, Jessen H, Gort SJ, Selmer T, Buckel W;

XX WPI: 2002-627219/57.

DR N-PSDB; ABS66661.

XX Novel polypeptides having lactyl-CoA dehydratase, E1 activator, 3-

PT hydroxypropionyl-CoA dehydratase or malonyl-CoA reductase activity,

PT useful in producing 3-hydroxypropionic acid and other organic compounds.

XX Claim 27; Fig 28; 237pp; English.

PS The invention describes an isolated polypeptide (I) useful for producing

XX 3-hydroxypropionic acid (3-HP), polymerised 3-HP, an ester of 3-HP,

CC polymerised acrylate, or an ester of acrylate. A cell having lactyl CoA-

CC dehydratase activity and 3-hydroxypropionyl-CoA dehydratase activity is  
 CC useful for producing 3-HP or an ester of 3-HP, or polymerised 3-HP. The 3  
 CC -HP produced is useful as food, feed or preservative. The nucleic acid  
 CC molecules are useful for engineering host cells with the ability to  
 CC produce 3-HP as well as other organic compounds such as 1,3-propane diol,  
 CC acrylic acid, polymerised acrylate or esters of acrylate. The  
 CC polypeptides can be used in cell-free systems to make 3-HP as well as  
 CC other organic compounds such as 1,3-propane diol, acrylic acid,  
 CC polymerised acrylate, esters of acrylate. The host cells can be used in  
 CC culture systems to produce large quantities of 3-HP as well as other  
 CC organic compounds as described above. This is the amino acid sequence of  
 CC a protein associated with the production of 3-hydroxypropionic acid

XX SQ Sequence 1822 AA;

Query Match

Best Local Similarity 87.9%; Score 29; DB 5; Length 1822;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRANLRA 7

DB 1554 RRANLRA 1560

Search completed: June 9, 2004, 14:57:08

Job time : 57 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2004, 14:55:11 ; Search time 20 Seconds  
(without alignments)  
33.667 Million cell updates/sec

Title: US-10-080-100-46

Perfect score: 33

Sequence: 1 Q9ANLRA 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	84.8	268	1	TSEBAT
2	28	84.8	268	2	AB0653
3	28	84.8	399	2	AD3512
4	28	84.8	483	2	G75392
5	28	84.8	502	2	A23547
6	27	81.8	269	2	C94932
7	27	81.8	318	2	T15869
8	27	81.8	390	2	AC3557
9	27	81.8	392	2	S11998
10	27	81.8	613	2	A82834
11	27	81.8	640	2	AG1269
12	27	81.8	640	2	A11631
13	27	81.8	666	2	D82511
14	27	81.8	1114	2	B86433
15	27	81.8	1358	2	S33633
16	26	78.8	106	2	T34560
17	26	78.8	153	2	BS6394
18	26	78.8	230	2	A70459
19	26	78.8	268	2	D82232
20	26	78.8	275	2	H83325
21	26	78.8	282	2	D59379
22	26	78.8	303	2	B85182
23	26	78.8	303	2	E98049
24	26	78.8	334	2	I40040
25	26	78.8	334	2	A44894
26	26	78.8	335	2	I40159
27	26	78.8	336	1	FLYB3
28	26	78.8	336	2	S70256
29	26	78.8	336	2	S70255

30 26 78.8 336 2 I40204  
31 26 78.8 336 2 I40041  
32 26 78.8 336 2 A38450  
33 26 78.8 336 2 S37728  
34 26 78.8 336 2 I40077  
35 26 78.8 336 2 I40135  
36 26 78.8 336 2 I40092  
37 26 78.8 336 2 I40079  
38 26 78.8 336 2 I40140  
39 26 78.8 336 2 I40088  
40 26 78.8 336 2 I40075  
41 26 78.8 336 2 I40134  
42 26 78.8 355 2 I40308  
43 26 78.8 360 2 T47170  
44 26 78.8 366 2 T42377  
45 26 78.8 382 2 AG0188

#### ALIGNMENTS

##### RESULT 1

##### TSEBAT

C:Species: Salmonella typhimurium

C:Date: 28-Feb-1980 #sequence\_revision 28-Feb-1980 #text\_change 24-Sep-1999

C:Accession: A93837; A92126; A01152

R:Nichols, B.P.; Yanofoisky, C.

Proc. Natl. Acad. Sci. U.S.A. 76, 5244-5248, 1979

A:Title: Nucleotide sequences of trpA of Salmonella typhimurium and Escherichia coli: ar

A:Reference number: A93837; MUID:80056671; PMID:388433

A:Accession: A93837

A:Molecule type: DNA

A:Residues: 1-268 <NIC>

A:Cross-references: GB:V01376; NID:G47938; PIDN:CAA24666.1; PID:G47940

R:Li, S.B.; Yanofoisky, C.

J. Biol. Chem. 248, 1830-1836, 1973

A:Title: Amino acid sequence studies with the tryptophan synthetase alpha chain of Salmc

A:Reference number: A92126; MUID:73149276; PMID:4571777

A:Contents: tentative sequence

A:Accession: A92126

A:Molecule type: Protein

A:Residues: 1-268 <LIS>

C:Genetics:

A:Gene: trpA

C:Complex: heterotetramer; two alpha and two beta chains

C:Function:

A:Description: catalyzes conversion of indoleglycerol phosphate and serine to tryptophar

A:Pathway: tryptophan biosynthesis

A>Note: Cofactor Pyridoxal phosphate

A>Note: last step in pathway

C:Superfamily: tryptophan synthase alpha chain; tryptophan synthase alpha chain homology

C:Keywords: carbon-oxygen lyase; heterotetramer; hydro-lyase; tryptophan biosynthesis

F:18-246/Domain: tryptophan synthase alpha chain homology <TRPA>

F:49/Active site: Glu #status predicted

Query Match 84.8%; Score 28; DB 1; Length 268;

Best Local Similarity 85.7%; Pred. No. 46;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 Q9ANLRA 7

DB 65 Q9ANLRA 71

##### RESULT 2

##### AB0653

C:Species: Salmonella enterica subsp. enterica serovar typhi

A>Note: this species has also been called Salmonella typhi

C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002

C:Accession: AB0653

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher

th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.  
 Nature 413, 848-852, 2001  
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.  
 A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov  
 A:Reference number: AB0502; MUID:21534947; PMID:11677608  
 A:Accession: AB0653  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-268 <PAR>  
 A:Cross-references: GB:AL513382; PIDN:CAD08405.1; PID:g16502448; GSPDB:GN00176  
 C:Genetics:  
 A:Gene: STY1324  
 C:Superfamily: tryptophan synthase alpha chain; tryptophan synthase alpha chain homology

Query Match 84.8%; Score 28; DB 2; Length 268;  
 Best Local Similarity 85.7%; Pred. No. 46;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QANLRA 7  
 |||||  
 Db 65 QANLRA 71

RESULT 3  
 A:Residues: 1-399 <KUR>  
 A:Cross-references: GB:AE008918; PIDN:AAL53263.1; PID:g17984144; GSPDB:GN00191  
 A:Experimental source: strain 16M  
 C:Genetics:  
 A:Gene: BMEI10022  
 A:Map position: II

Query Match 84.8%; Score 28; DB 2; Length 399;  
 Best Local Similarity 85.7%; Pred. No. 68;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QANLRA 7  
 |||||  
 Db 4 QANLRA 10

RESULT 4  
 glycosyl hydrolase, family 13 - *Deinococcus radiodurans* (strain R1)  
 C:Species: *Deinococcus radiodurans*  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000  
 A:Accession: G75392  
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma  
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999  
 A:Title: Genome sequence of the radioreistant bacterium *Deinococcus radiodurans* R1.  
 A:Reference number: A75250; MUID:20036896; PMID:10567266  
 A:Accession: G75392  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-483 <WHI>  
 A:Cross-references: GB:AE001991; GB:AE000513; NID:g6459223; PIDN:AAF11034.1; PID:g645922  
 A:Experimental source: strain R1  
 C:Genetics:

A:Gene: DRI472  
 A:Map position: 1  
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology

Query Match 84.8%; Score 28; DB 2; Length 483;  
 Best Local Similarity 100.0%; Pred. No. 82;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RANLRA 7  
 |||||  
 Db 54 RANLRA 59

RESULT 5  
 A:Residues: 1-502 <PRA>  
 A:Cross-references: GB:M13811; NID:g214555; PIDN:AAA49891.1; PID:g214556  
 C:Superfamily: cytoskeletal keratin  
 C:Keywords: coiled coil

Query Match 84.8%; Score 28; DB 2; Length 502;  
 Best Local Similarity 85.7%; Pred. No. 85;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QANLRA 7  
 |||||  
 Db 335 QANLRA 341

RESULT 6  
 tryptophan synthase (EC 4.2.1.20) alpha chain [imported] - *Buchnera* sp. (strain APS)  
 C:Species: *Buchnera* sp.  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 23-Mar-2001  
 C:Accession: C84962  
 R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.  
 Nature 407, 81-86, 2000  
 A:Title: Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. A  
 A:Reference number: A84930; MUID:20445173; PMID:10933077  
 A:Accession: C84962  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-269 <STO>  
 A:Cross-references: GB:AP000398; GSPDB:GN00144  
 A:Experimental source: strain APS  
 C:Genetics:  
 A:Gene: trpa; BU277  
 C:Superfamily: tryptophan synthase alpha chain; tryptophan synthase alpha chain homology  
 C:Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 81.8%; Score 27; DB 2; Length 269;  
 Best Local Similarity 71.4%; Pred. No. 79;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QANLRA 7  
 |||||  
 Db 65 QANLRA 71

RESULT 7  
 T15869  
 hypothetical protein C56G2.3 - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*

C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
C:Accession: T15869  
R:Connell, M.  
A:Description: The sequence of C. elegans cosmid C56G2.  
A:Reference number: Z18420  
A:Accession: T15869  
A:Molecule type: DNA  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Residues: 1-318 <CON>  
A:Cross-references: EMBL:U23177; NID:g726411; PID:g726415; PIDN:AAA64330.1; CESP:C56G2.3  
A:Experimental source: strain Bristol N2  
C:Genetics:  
A:Gene: CESP:C56G2.3  
A:Introns: 178/1

Query Match 81.8%; Score 27; DB 2; Length 318;  
Best Local Similarity 71.4%; Pred. No. 94;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRANLRA 7  
:|||||:  
Db 226 RRANIRA 232

RESULT 8  
AC3557  
acriflavin resistance protein a precursor [imported] - Brucella melitensis (strain 16M)  
C:Species: Brucella melitensis  
C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 01-Feb-2002  
C:Accession: AC3557  
R:DelVecchio, V.; Kapatal, R.J.; Redkar, R.J.; Patra, G.; Mufer, C.; Los, T.; Ivanova,  
.; Mazur, M.; Gotsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis  
A:Reference number: AD3252; PMID:11756688  
A:Accession: AC3557  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-390 <KUR>  
A:Cross-references: GB:AE008919; PIDN:AAL53622.1; PID:g17984537; GSPDB:GN00191  
A:Experimental source: strain 16M  
C:Genetics:  
A:Gene: BMEII0380  
A:Map position: II

Query Match 81.8%; Score 27; DB 2; Length 390;  
Best Local Similarity 83.3%; Pred. No. 1.2e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRANLR 6  
:|||||:  
Db 121 QRANVR 126

RESULT 9  
S11998  
finger protein odd-skipped - fruit fly (Drosophila melanogaster)  
C:Species: Drosophila melanogaster  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 24-Jul-1997  
C:Accession: S11998  
R:Coulter, D.E.; Swayuk, E.A.; Beran-Koehn, M.A.; Goldberg, D.; Wieschaus, E.; Schedl,  
EMBO J. 8, 3795-3804, 1990  
A:Title: Molecular analysis of odd-skipped, a zinc finger encoding segmentation gene with  
A:Reference number: S11998  
A:Accession: S11998  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-392 <COU>

Query Match 81.8%; Score 27; DB 2; Length 392;  
Best Local Similarity 71.4%; Pred. No. 1.2e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRANLRA 7  
:|||||:  
Db 315 QRANLKS 321

## RESULT 10

A82834  
hypothetical protein XF0221 [imported] - Xylella fastidiosa (strain 9a5c)  
C:Species: Xylella fastidiosa  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: A82834  
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen

Nature 406, 151-157, 2000  
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: A82515; MUID:20365717; PMID:10910347  
A>Note: for a complete list of authors see reference number A59328 below  
A:Accession: A82834  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-613 <SIM>  
A:Cross-references: GB:AE003875; GB:AE003849; NID:g9105019; PIDN:AAF83034.1; GSPDB:GN001  
A:Experimental source: strain 9a5c  
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A  
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier, H  
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.  
submitted to GenBank, June 2000  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig  
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins E  
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.P.M.; Miracca, E.C.; Miyaki, C.Y.  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak  
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir  
M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: XF0221

Query Match 81.8%; Score 27; DB 2; Length 613;  
Best Local Similarity 85.7%; Pred. No. 1.8e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QRANLRA 7  
:|||||:  
Db 240 QRAQLRA 246

## RESULT 11

AG1269  
threonyl-tRNA synthetase [imported] - Listeria monocytogenes (strain EGD-e)  
C:Species: Listeria monocytogenes  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 14-Dec-2001  
C:Accession: AG1269  
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
Science 294, 849-852, 2001  
A:Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma  
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,  
A:Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AG1269  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-640 <GLA>  
A:Cross-references: GB:NC 003210; PIDN:CAC99637.1; PID:g16410988; GSPDB:GN00177  
A:Experimental source: strain EGD-e  
C:Genetics:  
A:Gene: thrS  
C:Superfamily: threonine-tRNA ligase

Query Match 81.8%; Score 27; DB 2; Length 640;  
 Best Local Similarity 85.7%; Pred. No. 1.9e+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QRANLRA 7  
 |||||  
 Db 568 QRAGLRA 574

RESULT 12  
 A:Accession: A11631  
 C:Species: Listeria innocua  
 C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 14-Dec-2001  
 C:Accession: A11631  
 R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001  
 A:Authors: Krefte, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Meakins, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, W.; Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.  
 A:Title: Comparative Genomics of *Listeria* species.  
 A:Reference number: AB1077; MUID:21537279; PMID:11679669  
 A:Accession: A11631  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-640 <GLA>  
 A:Cross-references: GB:AL592022; PIDN:CA096825.1; PID:G16414081; GSPDB:GN00178  
 A:Experimental source: strain Clp11262  
 C:Genetics:  
 A:Gene: thrS  
 C:Superfamily: threonine-tRNA ligase

Query Match 81.8%; Score 27; DB 2; Length 640;  
 Best Local Similarity 85.7%; Pred. No. 1.9e+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QRANLRA 7  
 |||||  
 Db 568 QRAGLRA 574

RESULT 13  
 D82511  
 C:Species: Vibrio cholerae  
 C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
 C:Accession: D82511  
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, J.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Nature 406, 477-483, 2000  
 A:Title: DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.  
 A:Reference number: A82035; MUID:20406833; PMID:10952301  
 A:Accession: D82511  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-666 <HEI>  
 A:Cross-references: GB:AE004345; GB:AE003853; NID:G9657390; PIDN:AA095930.1; GSPDB:GN00178  
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
 C:Genetics:  
 A:Gene: VCA0016  
 A:Map position: 2  
 C:Superfamily: 1,4-alpha-glucan branching enzyme

Query Match 81.8%; Score 27; DB 2; Length 666;  
 Best Local Similarity 85.7%; Pred. No. 2e+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QRANLRA 7  
 |||||  
 Db 486 QRANLRA 492

RESULT 14  
 B86423  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Dec-2001  
 C:Accession: B86423  
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.B.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, R.; M. J.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.; Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
 A:Reference number: A86141; MUID:21016719; PMID:11130712  
 A:Accession: B86423  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1114 <STO>  
 A:Cross-references: GB:AE005172; NID:G10092454; PIDN:AAG12856.1; GSPDB:GN00141  
 C:Genetics:  
 A:Map position: 1

Query Match 81.8%; Score 27; DB 2; Length 1114;  
 Best Local Similarity 83.3%; Pred. No. 3.3e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRANLR 6  
 |||||  
 Db 817 QRANIR 822

RESULT 15  
 S33653  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 30-Sep-1993 #sequence\_revision 02-Aug-1994 #text\_change 24-Sep-1999  
 C:Accession: S33653; S36717; S36732; JH0486  
 R:Clark, M.W.; Zhong, W.W.; Keng, T.; Storms, R.K.; Ouellette, B.F.F.; Barton, A.; Kaback  
 Yeast 9, 543-549, 1993  
 A:Title: The YAL017 gene on the left arm of chromosome I of *Saccharomyces cerevisiae* encodes a probable serine/threonine protein kinase (EC 2.7.1.-).  
 A:Reference number: S33653; MUID:93311122; PMID:8322517  
 A:Accession: S33653  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-1358 <CLA>  
 A:Cross-references: EMBL:L05146  
 R:Ouellette, B.F.; Clark, M.W.; Keng, T.; Storms, R.K.; Zhong, W.; Zeng, B.; Fortin, N.; De  
 Submitted to the EMBL data library, January 1993  
 A:Description: Sequencing of Chromosome I from *Saccharomyces cerevisiae*: analysis of a 32  
 A:Reference number: S36717  
 A:Accession: S36717  
 A:Molecule type: DNA  
 A:Residues: 1-864,867-1358 <QUE>  
 A:Cross-references: EMBL:L05146; NID:G171851; PIDN:AA04940.1; PID:G171858; MIPS:YAL017W  
 R:Clark, M.W.; Zhong, W.W.; Keng, T.; Storms, R.K.; Barton, A.; Kaback, D.B.; Sussey, H.  
 Yeast 8, 133-145, 1992  
 A:Title: Identification of a *Saccharomyces cerevisiae* homolog of the SNF2 transcriptional  
 A:Reference number: S22266; MUID:92221690; PMID:1561836  
 A:Accession: S36732  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-862 <CL2>  
 A:Cross-references: EMBL:S93805  
 R:Sidhu, R.S.; Mathewes, S.; Bollon, A.P.  
 Gene 107, 111-118, 1991  
 A:Title: Selection of secretory protein-encoding genes by fusion with PHO5 in *Saccharomyces*  
 A:Reference number: JH0483; MUID:92077420; PMID:1743509  
 A:Accession: JH0486

A;Molecule type: DNA  
A;Residues: 1-72,'E',74-154 <SID>  
C;Genetics:  
A;Gene: SGD:FUN31; SSPI38  
A;Cross-references: SGD:S0000015; MIPS:YAL017w  
A;Map position: 1L  
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo  
C;Keywords: ATP; glycoprotein; phosphotransferase; serine/threonine-specific protein kin  
F;1096-1356/Domain: protein kinase homology <KIN>  
F;1104-1112/Region: protein kinase ATP-binding motif  
F;8.125/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;1232/Active site: Asp #status predicted

Query Match 81.8%; Score 27; DB 2; Length 1358;  
Best Local Similarity 71.4%; Pred. No. 4e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRANLRA 7  
Db 334 QRANLKS 340

Search completed: June 9, 2004, 14:58:58  
Job time : 21 secs



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2004, 14:51:20 ; Search time 11 Seconds  
(without alignments)  
33.136 Million cell updates/sec

Title: US-10-080-100-46

Perfect score: 33

Sequence: 1 Q9ANLRA 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	90.9	270	1 TRPA_BUCDN	O68429 buchnera ap
2	30	90.9	270	1 TRPA_BUCMH	Q9rq33 buchnera ap
3	29	87.9	269	1 TRPA_BUCBP	P59457 buchnera ap
4	28	84.8	268	1 TRPA_SALTI	Q827e0 salmonella
5	28	84.8	268	1 TRPA_SALTY	P00929 salmonella
6	28	84.8	502	1 K2C8_XENLA	P8776 xenopus lae
7	27	81.8	269	1 TRPA_BUCAI	P57365 buchnera ap
8	27	81.8	318	1 YQK3_CABEL	Q09287 caenorhabdi
9	27	81.8	332	1 ODD_DROME	P23803 grosophila
10	27	81.8	640	1 SYT_LISIN	Q92bf5 listeria in
11	27	81.8	640	1 SYT_LISMO	Q8y6x2 listeria mo
12	27	81.8	666	1 GLGB_VIBCH	Q9kne8 vibrio chol
13	27	81.8	715	1 GLGB_VIBVU	Q8d4p0 vibrio vuln
14	27	81.8	755	1 GLGB_VIBPA	Q87fr0 vibrio para
15	27	81.8	1356	1 KAB7_YEAST	P31374 saccharomyc
16	26	78.8	230	1 YI45_AQUAE	Q67699 aquifex aeo
17	26	78.8	268	1 TRPA_VIBCH	Q9kst7 vibrio chol
18	26	78.8	275	1 TAM_PSBAB	Q91os1 pseudomonas
19	26	78.8	310	1 Y4RE_RHISN	P55338 rhizobium s
20	26	78.8	336	1 FLAI_BORBU	P11089 borrelia bu
21	26	78.8	355	1 YGR1_SCHPO	Q9uuc2 schizosacch
22	26	78.8	382	1 HIS8_YERPE	Q8zic6 versinia pe
23	26	78.8	382	1 Y894_HAEIN	Q57500 haemophilus
24	26	78.8	402	1 C13B_XYLFA	Q9p5c5 xyloella fas
25	26	78.8	402	1 C13B_XYLFT	Q87ax5 xyloella fas
26	26	78.8	459	1 EX7L_YERPE	Q8zcu2 versinia pe
27	26	78.8	467	1 SES4_YEAST	Q02825 saccharomyc
28	26	78.8	514	1 MALF_ECOLI	P02916 escherichia
29	26	78.8	514	1 MALF_ENTAE	P18812 enterobacte
30	26	78.8	514	1 MALF_SALTY	P26467 salmonella
31	26	78.8	514	1 MALF_SHIFL	Q83p81 shigella fl
32	26	78.8	602	1 YH02_YEAST	P38887 saccharomyc
33	26	78.8	626	1 CC23_YEAST	P16522 saccharomyc

#### ALIGNMENTS

RESULT 1

```

TRPA_BUCDN
ID TRPA_BUCDN STANDARD; PRT; 270 AA.
AC O68429;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tryptophan synthase alpha chain (EC 4.2.1.20).
GN TRPA.
OS Buchnera aphidicola (subsp. Diuraphis noxia).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=118101;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98290710; PubMed=9625791;
RA Baumann L., Baumann P., Moran N.A.;
RT "The endosymbiont (Buchnera) of the aphid Diuraphis noxia contains all
RL Curr. Microbiol. 37:58-59(1998).
CC -!- FUNCTION: The alpha subunit is responsible for the aldol cleavage
CC of indoleglycerol phosphate to indole and glyceraldehyde 3-
CC phosphate.
CC -!- CATALYTIC ACTIVITY: L-serine + 1-(indol-3-yl)glycerol 3-phosphate
CC = L-tryptophan + glyceraldehyde 3-phosphate.
CC -!- PATHWAY: Tryptophan biosynthesis; fifth (last) step.
CC -!- SUBUNIT: Tetramer of two alpha and two beta chains (By
CC similarity).
CC -!- SIMILARITY: Belongs to the trpA family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF038565; AAC27736.1; -.
CC HSSP; P00929; 2MSY.
CC HAMAP; MF_00131; -.
CC InterPro; IPR003009; FMN enzyme.
CC InterPro; IPR002028; Trp_synthaseA.
CC Pfam; PF00290; trp_synA; 1.
CC ProDom; PD001535; Trp_synthaseA; 1.
CC TIGRPFAMs; TIGR00262; trpA; 1.
CC PROSITE; PS00167; TRP_SYNTHASE_ALPHA; 1.
CC Tryptophan biosynthesis; Lyase.
CC SEQUENCE 270 AA; 30442 MW; D8D3E73C4CDB1CCD CRC64;

```

Query Match 90.9%; Score 30; DB 1; Length 270;  
Best Local Similarity 85.7%; Pred. No. 6;  
Matches 6; Conservative 1; Mismatches 0; Gaps 0;

QY 1 Q9ANLRA 7

|||||

```

Db          65 QKANLRA 71
RESULT 2
TRPA_BUCBP
ID TRPA_BUCBP STANDARD; PRT; 270 AA.
AC Q9RQ33; 2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tryptophan synthase alpha chain (EC 4.2.1.20).
GN TRPA
OS Buchnera aphidicola (subsp. Melaphis rhois).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=118103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20022990; PubMed=10555290;
RA Clark M.A., Moran N.A., Baumann P.;
RT "Sequence evolution in bacterial endosymbionts having extreme base
compositions.";
RL Mol. Biol. Evol. 16:1586-1598(1999).
CC -!- FUNCTION: The alpha subunit is responsible for the aldol cleavage
of indoleglycerol phosphate to indole and glyceraldehyde 3-
phosphate.
CC -!- CATALYTIC ACTIVITY: L-serine + 1-(indol-3-yl)glycerol 3-phosphate
= L-tryptophan + glyceraldehyde 3-phosphate.
CC -!- PATHWAY: Tryptophan biosynthesis; fifth (last) step.
CC -!- SUBUNIT: Tetramer of two alpha and two beta chains (By
similarity).
CC -!- SIMILARITY: Belongs to the trpA family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; AF132318; AAF14255.1; -.
CC HSSP; P00929; 2WSY.
CC HAMAP; MF 00131; -.
CC InterPro; IPR003009; FMN enzyme.
CC InterPro; IPR002028; Trp_synthaseA.
CC Pfam; PF00290; trp_synthA; 1.
CC ProDom; PD001535; Trp_synthaseA; 1.
CC TIGRfams; TIGR00262; trpA; 1.
CC PROSITE; PS00167; TRP SYNTHASE ALPHA; 1.
CC Tryptophan biosynthesis; Lyase.
CC KW Tryptophan biosynthesis; Lyase.
CC SEQUENCE 270 AA; 30398 MW; 9FBECE911836C3EC CRC64;
SQ
Query Match 90.9%; Score 30; DB 1; Length 270;
Best Local Similarity 85.7%; Pred. No. 6;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 QKANLRA 7
|:|||||
Db 65 QKANLRA 71
RESULT 3
TRPA_BUCBP
ID TRPA_BUCBP STANDARD; PRT; 269 AA.
AC P59457;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tryptophan synthase alpha chain (EC 4.2.1.20).
GN TRPA OR BB257.
OS Buchnera aphidicola (subsp. Baizongia pistaciae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

```

```

OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=135842;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22456901; PubMed=12522265;
RA Van Ham R.C.H.J., Kamerbeek J., Palacios C., Rausell C., Abascal F.,
Bastolla U., Fernandez J.M., Jimenez L., Postigo M., Silva F.O.,
Tamames J., Viguera E., Latorre A., Valencia A., Moran F., Moya A.;
RT "Reductive genome evolution in Buchnera aphidicola.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:581-586(2003).
CC -!- FUNCTION: The alpha subunit is responsible for the aldol cleavage
of indoleglycerol phosphate to indole and glyceraldehyde 3-
phosphate.
CC -!- CATALYTIC ACTIVITY: L-serine + 1-(indol-3-yl)glycerol 3-phosphate
= L-tryptophan + glyceraldehyde 3-phosphate.
CC -!- PATHWAY: Tryptophan biosynthesis; fifth (last) step.
CC -!- SUBUNIT: Tetramer of two alpha and two beta chains (By
similarity).
CC -!- SIMILARITY: Belongs to the trpA family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; AE014016; AA026984.1; -.
CC HAMAP; MF 00131; -.
CC InterPro; IPR003009; FMN enzyme.
CC InterPro; IPR002028; Trp_synthaseA.
CC Pfam; PF00290; trp_synthA; 1.
CC TIGRfams; TIGR00262; trpA; 1.
CC PROSITE; PS00167; TRP SYNTHASE ALPHA; 1.
CC Tryptophan biosynthesis; Lyase; Complete proteome.
CC KW Tryptophan biosynthesis; Lyase.
CC SEQUENCE 269 AA; 30078 MW; 42E4AA359B8162E8 CRC64;
SQ
Query Match 87.9%; Score 29; DB 1; Length 269;
Best Local Similarity 85.7%; Pred. No. 11;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 QKANLRA 7
|:|||||
Db 66 QKANLRA 72
RESULT 4
TRPA_SALTI
ID TRPA_SALTI STANDARD; PRT; 268 AA.
AC Q9Z7E0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tryptophan synthase alpha chain (EC 4.2.1.20).
GN TRPA OR STY1324 OR T1639.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Bentley S.D., Holden M.T.G., Sebahia M.,
Churcher C., Mungall K.L., Bentley S.D., Chillingworth T., Connor P.,
Baker S., Basham D., Brooks K., Davies R.M., Dowd L., White N., Farrar J.,
Cronin A., Davis P., Davies R.M., Hien T.T., Holroyd S., Jagels K.,
Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
Krogh A., Larsen T.S., Leather S., Moule S., O'Garra P., Parry C.,
Raoult M.A., Rutherford K., Simmonds M., Skelton J., Stevens K.,
Whitehead S., Barrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella

```



RT enterica serovar Typhi CT18.";  
RL Nature 413:848-852(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Ty2 / ATCC 700931;  
RX MEDLINE=22531367; PubMed=12644504;  
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,  
Burland V., Kodoyianni V., Schwartz D.C., Blattner P.R.,  
"Comparative genomics of Salmonella enterica serovar Typhi strains Ty2  
and CT18.";  
RT J. Bacteriol. 185:2330-2337(2003).  
CC -!- FUNCTION: The alpha subunit is responsible for the aldol cleavage  
of indoleglycerol phosphate to indole and glyceraldehyde 3-  
phosphate.  
CC -!- CATALYTIC ACTIVITY: L-serine + 1-(indol-3-yl)glycerol 3-phosphate  
= L-tryptophan + glyceraldehyde 3-phosphate.  
CC -!- PATHWAY: Tryptophan biosynthesis; fifth (last) step.  
CC -!- SUBUNIT: Tetramer of two alpha and two beta chains (By  
similarity).  
CC -!- SIMILARITY: Belongs to the trpA family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
at the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AL627269; CAD08405.1; -;  
DR EMBL; A5016839; AAC69266.1; -;  
DR HAMAP; MF\_00131; -; 1.  
DR InterPro; IPR003009; FMN enzyme.  
DR InterPro; IPR002028; Trp synthase.  
DR Pfam; PF00290; trp syntA; 1.  
DR ProDom; PD001535; Trp synthase; 1.  
DR TIGRFAMs; TIGR00262; trpA; 1.  
DR PROSITE; PS00167; TRP SYNTHASE ALPHA; 1.  
KW Tryptophan biosynthesis; Lyase; Complete proteome.  
SQ SEQUENCE 268 AA; 28671 MW; F409BF1093B581B5 CRC64;

Query Match 84.8%; Score 28; DB 1; Length 268;  
Best Local Similarity 85.7%; Pred.No.18;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QNANLEA 7  
| | | | |  
Db 65 QNANLEA 71

RESULT 5  
ID TRPA\_SALTY STANDARD; PRT; 268 AA.  
AC F00929;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Tryptophan synthase alpha chain (EC 4.2.1.20).  
GN TRPA OR STM1727.  
OS Salmonella typhimurium.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Salmonella.  
OX NCBI\_TaxID=602;  
RN [1]\_TaxID=602;  
RP SEQUENCE FROM N.A.  
RX MEDLINE=80056671; PubMed=388433;  
RA Nichols B.P., Yanofsky C.;  
RT "Nucleotide sequences of trpA of Salmonella typhimurium and  
Escherichia coli: an evolutionary comparison.";  
RL Proc. Natl. Acad. Sci. U.S.A. 76:5244-5248(1979).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=81223839; PubMed=7017727;

RA Schneider W.P., Nichols B.P., Yanofsky C.;  
RT "Procedure for production of hybrid genes and proteins and its use in  
assessing significance of amino acid differences in homologous  
tryptophan synthetase alpha polypeptides.";  
RL Proc. Natl. Acad. Sci. U.S.A. 78:2169-2173(1981).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;  
RX MEDLINE=21534948; PubMed=11677609;  
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,  
Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,  
Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,  
Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,  
Waterston R., Wilson R.K.;  
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium  
LT2.";  
RL Nature 413:852-856(2001).  
RN [4]  
RP PARTIAL SEQUENCE.  
RX MEDLINE=73149276; PubMed=4571777;  
RA Li S.-L., Yanofsky C.;  
RT "Amino acid sequence studies with the tryptophan synthetase alpha  
chain of Salmonella typhimurium.";  
RL J. Biol. Chem. 248:1830-1836(1973).  
RN [5]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).  
RX MEDLINE=89034326; PubMed=3053720;  
RA Hyde C.C., Ahmed S.A., Padlan E.A., Miles E.W., Davies D.R.;  
RT "Three-dimensional structure of the tryptophan synthase alpha 2 beta  
2 multienzyme complex from Salmonella typhimurium.";  
RL J. Biol. Chem. 263:17857-17871(1988).  
RN [6]  
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).  
RX MEDLINE=97352620; PubMed=9201907;  
RA Rhee S., Parris K.D., Hyde C.C., Ahmed S.A., Miles E.W., Davies D.R.;  
RT "Crystal structures of a mutant (betak87T) tryptophan synthase  
alpha2beta2 complex with ligands bound to the active sites of the  
alpha- and beta-subunits reveal ligand-induced conformational  
changes.";  
RL Biochemistry 36:7664-7680(1997).  
RN [7]  
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).  
RX MEDLINE=98204834; PubMed=9535826;  
RA Rhee S., Miles E.W., Davies D.R.;  
RT "Cryo-crystallography of a true substrate, indole-3-glycerol  
phosphate, bound to a mutant (alpha60N) tryptophan synthase  
alpha2beta2 complex reveals the correct orientation of active site  
alphaGlu49.";  
RL J. Biol. Chem. 273:8553-8555(1998).  
RN [8]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RA Hyde C.C., Parris K.D., Bhat T.N., Brown C., Ahmed S.A., Miles E.W.,  
Davies D.R.;  
RT "Refined structure of the native form of the tryptophan synthase  
multienzyme complex from Salmonella typhimurium.";  
RL Submitted (JUL-1998) to the PDB data bank.  
RN [9]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
RX MEDLINE=99435740; PubMed=10504236;  
RA Sachpatzidis A., Dealwis C., Lubetsky J.B., Liang P.-H.,  
Anderson K.S., Lolis E.;  
RT "Crystallographic studies of phosphonate-based alpha-reaction  
transition-state analogues complexed to tryptophan synthase.";  
RL Biochemistry 38:12665-12674(1999).  
RN [10]  
RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS).  
RX MEDLINE=40576265; PubMed=11034989;  
RA Weyand M., Schlichting I.;  
RT "Structural basis for the impaired channeling and allosteric  
inter-subunit communication in the beta A169L/beta C170W mutant of  
tryptophan synthase.";  
RL J. Biol. Chem. 275:41058-41063(2000).  
CC -!- FUNCTION: The alpha subunit is responsible for the aldol cleavage

CC of indoleglycerol phosphate to indole and glyceraldehyde 3-phosphate

CC -!- CATALYTIC ACTIVITY: L-serine + 1-(indol-3-yl)glycerol 3-phosphate

CC = L-tryptophan + glyceraldehyde 3-phosphate.

CC -!- PATHWAY: tryptophan biosynthesis; fifth (last) step.

CC -!- SUBUNIT: Tetramer of two alpha and two beta chains.

CC -!- SIMILARITY: Belongs to the trpA family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; V01376; CAA24666.1; -

CC EMBL; J01810; AAA27235.1; -

CC EMBL; AE008776; AAL20645.1; -

CC PIR; A93837; TSEBAT.

CC PDB; 2WSY; 30-MAR-99.

CC PDB; 1TTP; 08-MAR-96.

CC PDB; 1TQO; 08-MAR-96.

CC PDB; 2TRS; 01-APR-97.

CC PDB; 2TSI; 01-APR-97.

CC PDB; 2TYS; 01-APR-97.

CC PDB; 1UBS; 08-MAR-96.

CC PDB; 1A5A; 27-MAY-98.

CC PDB; 1A5B; 27-MAY-98.

CC PDB; 1A5S; 30-MAR-99.

CC PDB; 1A50; 30-MAR-99.

CC PDB; 1BEU; 12-AUG-98.

CC PDB; 1BKS; 23-MAR-99.

CC PDB; 1C29; 14-JAN-00.

CC PDB; 1C8V; 13-JAN-00.

CC PDB; 1C9D; 29-DEC-99.

CC PDB; 1CW2; 29-DEC-99.

CC PDB; 1CX9; 10-JAN-00.

CC PDB; 1FUY; 17-JAN-01.

CC PDB; 1K3U; 03-JUL-02.

CC PDB; 1K7E; 10-JUL-02.

CC PDB; 1K7F; 10-JUL-02.

CC PDB; 1K7X; 19-JUN-02.

CC PDB; 1K8X; 18-DEC-02.

CC PDB; 1K8Y; 19-JUN-02.

CC PDB; 1K8Z; 19-JUN-02.

CC PDB; 1KFB; 07-JAN-03.

CC PDB; 1KFC; 07-JAN-03.

CC PDB; 1KFE; 07-JAN-03.

CC PDB; 1LOP; 10-NOV-00.

CC PDB; 1LOQ; 10-NOV-00.

CC StyGene; SGI0396; trpA.

CC HAMAP; MF\_00131; -; 1.

CC InterPro; IPR003009; FMN enzyme.

CC InterPro; IPR002028; Trp synthaseA.

CC Pfam; PF00290; trp syntA; 1.

CC ProDom; PD001535; Trp synthaseA; 1.

CC TIGRFAMs; TIGR00262; trpA; 1.

CC PROSITE; P500167; TRP SYNTHASE ALPHA; 1.

CC tryptophan biosynthesis; Lyase; 3D-structure; Complete proteome.

CC FT HELIX 2 13

CC FT TURN 14 15

CC FT STRAND 18 24

CC FT TURN 25 26

CC FT HELIX 30 43

CC FT TURN 44 44

CC FT STRAND 48 51

CC FT TURN 58 59

CC FT HELIX 62 73

CC FT TURN 74 75

CC FT HELIX 78 91

CC FT STRAND 97 101

CC FT HELIX 103 107

FT TURN 108 109

FT HELIX 111 121

FT TURN 122 122

FT STRAND 125 128

FT TURN 129 130

FT HELIX 133 135

FT HELIX 137 145

FT TURN 146 147

FT STRAND 149 151

FT STRAND 153 154

FT TURN 156 157

FT HELIX 160 169

FT STRAND 174 176

FT HELIX 193 203

FT TURN 204 204

FT STRAND 208 210

FT HELIX 217 226

FT TURN 227 227

FT STRAND 230 233

FT HELIX 236 243

FT TURN 244 246

FT HELIX 248 264

FT HELIX 265 268

SQ SEQUENCE 268 AA; 25670 MW; F409BF1A931581B5 CRC64;

Query Match 84.8%; Score 28; DB 1; Length 268;

Best Local Similarity 85.7%; Pred. NO. 18;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QANLRA 7

Db 65 QANLRA 71

-----

RESULT 6

K2C8 XENLA STANDARD; PRT; 502 AA.

AC P08776;

DT 01-NOV-1988 (Rel. 09, Created)

DT 01-NOV-1988 (Rel. 09, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Keratin, type II cytoskeletal 8 (Cytokeratin 8).

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;

OC Xenopodinae; Xenopus.

OX NCBI\_TaxID=8355;

RI [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=86313601; PubMed=2428034;

RA Franz J.K., Franke W.W.;

RT "Cloning of cDNA and amino acid sequence of a cytokeratin expressed

RL in oocytes of Xenopus laevis."

RL Proc. Natl. Acad. Sci. U.S.A. 83:6475-6479 (1986).

CC -!- SUBUNIT: HETEROtetramer of two type I and two type II KERATINS.

CC KERATIN 8 ASSOCIATES WITH KERATIN 18.

CC -!- DEVELOPMENTAL STAGE: SYNTHESIZED IN THE OOCYTE IN EARLY AND LATE

CC EMBRYONIC STAGES.

CC -!- MISCELLANEOUS: THERE ARE TWO TYPES OF CYTOSKELETAL AND

CC MICROFILLLAR KERATIN, I (ACIDIC) AND II (NEUTRAL TO BASIC)

CC (40-55 AND 56-70 KILODALTONS, RESPECTIVELY).

CC -!- SIMILARITY: Belongs to the intermediate filament family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; M13811; AAA49891.1; -

CC PIR; A23547; A23547.

```

DR InterPro; IPR001664; IF.
DR InterPro; IPR002957; Keratin_I.
DR InterPro; IPR003054; Keratin_II.
DR Pfam; PF00038; filament; 1.
DR PRINTS; PR01248; TYPE1KERATIN.
DR PRINTS; PR01276; TYPE2KERATIN.
DR PROSITE; PS00226; IF; 1.
KW Intermediate filament; Coiled coil; Keratin; Phosphorylation.
FT DOMAIN 1 98 HEAD.
FT DOMAIN 99 406 ROD.
FT DOMAIN 407 502 TAIL.
FT DOMAIN 99 134 COIL 1A.
FT DOMAIN 135 151 LINKER 1.
FT DOMAIN 152 243 COIL 1B.
FT DOMAIN 244 267 LINKER 12.
FT DOMAIN 268 406 COIL 2.
FT SITE 350 350 STUTTER.
FT MOD_RES 26 26 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 502 AA; 55679 MW; 96249824FBE131F0 CRC64;

Query Match 84.8%; Score 28; DB 1; Length 502;
Best Local Similarity 85.7%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QRANLRA 7
DB 335 QRANLEA 341

RESULT 7
TRPA_BUCAI STANDARD; PRT; 269 AA.
AC P57365;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Trypophan synthase alpha chain (EC 4.2.1.20).
GN TRPA OR BU27.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum symbiotic bacterium).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=118099;
[1]
SEQUENCE FROM N.A.
RP STRAIN=Tokyo 1998;
RX MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. APS."
RT Buchnera sp. APS.
RL Nature 407:81-86(2000).
CC -!- FUNCTION: The alpha subunit is responsible for the aldol cleavage of indoleglycerol phosphate to indole and glyceraldehyde 3-phosphate.
CC -!- CATALYTIC ACTIVITY: L-serine + 1-(indol-3-yl)glycerol 3-phosphate = L-tryptophan + glyceraldehyde 3-phosphate.
CC -!- PATHWAY: Tryptophan biosynthesis; fifth (last) step.
CC -!- SUBUNIT: Tetramer of two alpha and two beta chains (By similarity).
CC -!- SIMILARITY: Belongs to the trpA family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC
CC EMBL; AF001118; BAB12987.1; -.
CC HSSP; P00929; 2WSY.
CC HAMAP; MF_00131; -. 1.
DR InterPro; IPR003009; FMN_enzyme.

```

```

DR InterPro; IPR002028; Trp synthaseA.
DR Pfam; PF00290; trp_syntA_1.
DR ProDom; PD001535; Trp_synthaseA; 1.
DR TIGRFAMs; TIGR00262; trpA; 1.
DR PROSITE; PS00167; TRP SYNTHASE ALPHA; 1.
KW Tryptophan biosynthesis; Lyase; Complete proteome.
SQ SEQUENCE 269 AA; 30322 MW; B1A59668F09BF3A5 CRC64;

Query Match 81.8%; Score 27; DB 1; Length 269;
Best Local Similarity 71.4%; Pred. No. 32;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRANLRA 7
DB 65 QXSNLRA 71

RESULT 8
YOK3_CABEL STANDARD; PRT; 318 AA.
ID YOK3_CABEL
AC Q09287;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 36.5 kDa protein C56G2.3 in chromosome III.
GN C56G2.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RA Connell M.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC
CC EMBL; U23177; AAA64330.1; -.
DR PIR; T15869; T15869.
DR WormPep; C56G2.3; CE01872.
DR InterPro; IPR007356; DUF425.
DR Pfam; PF04243; DUF425; 1.
KW Hypothetical protein.
SQ SEQUENCE 318 AA; 36521 MW; E4A8C962202DAB46 CRC64;

Query Match 81.8%; Score 27; DB 1; Length 318;
Best Local Similarity 71.4%; Pred. No. 39;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRANLRA 7
DB 226 RRANIRA 232

RESULT 9
ODD_DROME STANDARD; PRT; 392 AA.
ID ODD_DROME
AC P23803;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Odd-skipped protein.
GN ODD.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

```

OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91006082; PubMed=2120051;  
 RA Coulter D.E., Swaykay E.A., Beran-Koehn M.A., Goldberg D.,  
 RA Wieschaus E., Schedl P.;  
 RT "Molecular analysis of odd-skipped, a zinc finger encoding  
 RT segmentation gene with a novel pair-rule expression pattern.";  
 RL EMBO J. 9:3795-3804(1990).  
 CC -!- FUNCTION: PAIR-RULE PROTEIN THAT ESTABLISH PORTION OF ALTERNATING  
 CC SEGMENTS DURING EMBRYOGENESIS.  
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -!- SIMILARITY: Contains 4 C2H2-type zinc fingers.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X57480; CAA40718.1; -.  
 DR HSSP; P07248; 2ADR.  
 DR TRANSFAC; T00667; -.  
 DR FlyBase; FBgn002985; odd.  
 DR InterPro; IPR007087; Znf.C2H2.  
 DR Pfam; PF00096; zf-C2H2; 4.  
 DR ProDom; PD000003; Znf.C2H2; 1.  
 DR SMART; SM00355; Znf.C2H2; 4.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 4.  
 DR PROSITE; PS00157; ZINC\_FINGER\_C2H2\_2; 4.  
 DR Developmental protein; Pair-rule protein; Zinc-finger; Metal-binding;  
 KW DNA-binding; Repeat; Nuclear protein.  
 FT DOMAIN 82 93 POLY-GLN.  
 FT DOMAIN 101 106 POLY-GLN.  
 FT DOMAIN 142 149 POLY-GLN.  
 FT DOMAIN 150 181 HIS/PRO-RICH.  
 FT ZN FING 220 242 C2H2-TYPE 1.  
 FT ZN FING 248 270 C2H2-TYPE 2.  
 FT ZN FING 276 298 C2H2-TYPE 3.  
 FT ZN FING 304 326 C2H2-TYPE 4.  
 FT DOMAIN 369 375 POLY-SER.  
 FT SEQUENCE 392 AA; 44654 MW; 7272DD5C367F20B0 CRC64;  
 Query Match 81.8%; Score 27; DB 1; Length 392;  
 Best Local Similarity 71.4%; Pred. No. 48;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QRANLRA 7  
 Db 315 QRANLKS 321  
 RESULT 10  
 SYT\_LISMO STANDARD; PRT; 640 AA.  
 AC Q92EF5;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Threonyl-tRNA synthetase (EC 6.1.1.3) (Threonine--tRNA ligase)  
 DE (ThRS).  
 GN THRS OR LIN1594.  
 OS Listeria innocua.  
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
 OX NCBI\_TaxID=1642;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CLIP 11262 / Serovar 6a;  
 RX MEDLINE=21537279; PubMed=11679669;  
 RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,

BAQUERO F., Berche P., Bloecker H., Brandt P., Chakraborty T.,  
 Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,  
 Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,  
 Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,  
 Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,  
 Jones L.-M., Kaerst U., Krest J., Kuhn M., Kurst F., Kurapkat G.,  
 Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,  
 Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,  
 Remmel B., Rose M., Schluster T., Simoes N., Tierrez A.,  
 Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,  
 RT "Comparative genomics of *Listeria species*.";  
 RL Science 294:849-852(2001).  
 CC -!- CATALYTIC ACTIVITY: ATP + L-threonine + tRNA(Thr) = AMP +  
 CC diphosphate + L-threonyl-tRNA(Thr).  
 CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
 CC -!- SUBUNIT: Homodimer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AL596169; CAC96825.1; -.  
 DR PIR; A11631; A11631.  
 DR ListList; LIN01594; -.  
 DR HAMAP; MF\_00184; -; 1.  
 DR InterPro; IPR004154; HGTP anticodon.  
 DR InterPro; IPR004095; TGS dom.  
 DR InterPro; IPR002314; tRNA-synt 2b.  
 DR InterPro; IPR002320; tRNA-synt thr.  
 DR InterPro; IPR006195; tRNA ligase II.  
 DR Pfam; PF03129; HGTP anticodon; 1.  
 DR Pfam; PF02824; TGS; 1.  
 DR Pfam; PF00587; tRNA-synt 2b; 1.  
 DR PRINTS; PR01047; TRNASYNTHRR.  
 DR TIGRFAMS; TIGR00418; thrS; 1.  
 DR PROSITE; PS00862; AA tRNA LIGASE II; 1.  
 DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
 KW Metal-binding; Zinc; Complete proteome.  
 FT DOMAIN 241 538 CATALYTIC.  
 FT METAL 334 334 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT METAL 385 385 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT METAL 515 515 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT SEQUENCE 640 AA; 73209 MW; E42075466ADDF22 CRC64;  
 Query Match 81.8%; Score 27; DB 1; Length 640;  
 Best Local Similarity 85.7%; Pred. No. 82;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 QRANLRA 7  
 Db 568 QRAGLRA 574  
 RESULT 11  
 SYT\_LISMO STANDARD; PRT; 640 AA.  
 AC Q9V6X2;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Threonyl-tRNA synthetase (EC 6.1.1.3) (Threonine--tRNA ligase)  
 DE (ThRS).  
 GN THRS OR LMO1559.  
 OS Listeria monocytogenes.  
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
 OX NCBI\_TaxID=1639;  
 RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN-EGD-e / Serovar 1/2a;  
RX MEDLINE=21537279; PubMed=11679669;  
RA Glaser P., Frangoul L., Buchrieser C., Rusniok C., Amend A.,  
RA Baquero F., Berche P., Blocker H., Brandt P., Chakraborty T.,  
RA Charbit A., Chetouani F., Couve E., de Paruvier A., Deboux P.,  
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,  
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,  
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,  
RA Jones L.-M., Kaerst U., Krest J., Kuhn M., Kunst F., Kurapkat G.,  
RA Madueno E., Maitournan A., Mata Vicente J., Ng E., Nedjari H.,  
RA Nordisiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,  
RA Remmel B., Rose M., Schluefer T., Simoes N., Tierrez A.,  
RA Vazquez-Bolland J.-A., Voss H., Weiland J., Cossart P.,  
RT "Comparative genomics of *Listeria species*."  
RL Science 294:849-852(2001).  
CC -!- CATALYTIC ACTIVITY: ATP + L-threonine + tRNA(Thr) = AMP +  
CC diphosphate + L-threonyl-tRNA(Thr).  
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
CC -!- SUBUNIT: Homodimer (By similarity).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AL591979; CAC99637.1; -.  
CC F1; AG1269; AG1269.  
CC ListList; LMO01559; -.  
CC HAMAP; MF\_00184; -; 1.  
CC InterPro; IPR004154; HGTP anticodon.  
CC InterPro; IPR004095; TGS dom.  
CC InterPro; IPR002314; tRNA-synt 2b.  
CC InterPro; IPR002320; tRNA-synt thr.  
CC InterPro; IPR006195; tRNA ligase II.  
CC Pfam; PF03129; HGTP anticodon; 1.  
CC Pfam; PF02824; TGS; 1.  
CC Pfam; PF00587; tRNA-synt 2b; 1.  
CC PRINTS; PR01047; TRNASYNTHTR.  
CC TIGR; TIGR00418; thrS; 1.  
CC PROSITE; PS50822; AA\_TRNA\_LIGASE II; 1.  
CC Aminoacyl-tRNA synthetase; protein biosynthesis; Ligase; ATP-binding;  
CC Metal-binding; Zinc; Complete proteome.  
CC KW DOMAIN 241 538 CATALYTIC.  
CC FT METAL 334 334 ZINC (CATALYTIC) (BY SIMILARITY).  
CC FT METAL 385 385 ZINC (CATALYTIC) (BY SIMILARITY).  
CC FT METAL 515 515 ZINC (CATALYTIC) (BY SIMILARITY).  
CC SEQUENCE 640 AA; 73240 MW; 5B90914A14F94D30 CRC64;  
  
Query Match 81.8%; Score 27; DB 1; Length 640;  
Best Local Similarity 85.7%; Pred. No. 82;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ORANLRA 7  
DB 568 ORAGLRA 574  
  
RESULT 12  
GLGB\_VIBCH STANDARD; PRT; 666 AA.  
ID -GLGB\_VIBCH  
AC Q9KNE8;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE 1,4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching  
DE enzyme) (BE) (1,4-alpha-D-glucan:1,4-alpha-D-glucan 6-glucosyl-  
DE transferase).

GN GLGB OR VCA0016.  
OS Vibrio cholerae.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=666;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=El Tor N16961 / Serotype O1;  
RX MEDLINE=20406893; PubMed=10952301;  
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,  
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,  
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
RA Emolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,  
RA McDonald L., Uterback T., Fleischmann R.D., Nierman W.C., White O.,  
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
RA Fraser C.W.;  
RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio*  
RT cholerae."  
RL Nature 406:477-483(2000).  
CC -!- FUNCTION: Catalyzes the formation of the alpha-1,6-glucosidic  
CC linkages in glycogen by scission of a 1,4-alpha-linked  
CC oligosaccharide from growing alpha-1,4-glucan chains and the  
CC subsequent attachment of the oligosaccharide to the alpha-1,6  
CC position (By similarity).  
CC -!- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of  
CC glycogen.  
CC -!- PATHWAY: Glycogen biosynthesis; third step.  
CC -!- SUBUNIT: Monomer (By similarity).  
CC -!- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AE004345; AAF95930.1; -.  
CC F1; D82511; D82511.  
CC TIGR; VCA0016; -.  
CC HAMAP; MF\_00685; atypical; 1.  
CC InterPro; IPR006047; Alpha\_amyl\_cat.  
CC InterPro; IPR006407; GIGB.  
CC InterPro; IPR004193; Glyco\_hydro\_13N.  
CC Pfam; PF00128; alpha-amylase; 1.  
CC Pfam; PF02922; isoamylase N; 1.  
CC TIGR; TIGR01515; branching enzym; 1.  
CC Glycogen biosynthesis; Transferase; Glycosyltransferase;  
CC Complete proteome.  
CC KW ACT SITE 240 240 BY SIMILARITY.  
CC FT ACT SITE 240 240 BY SIMILARITY.  
CC FT ACT SITE 275 275 BY SIMILARITY.  
CC FT ACT SITE 280 280 BY SIMILARITY.  
CC FT ACT SITE 343 343 BY SIMILARITY.  
CC FT ACT SITE 345 345 BY SIMILARITY.  
CC FT ACT SITE 398 398 BY SIMILARITY.  
CC FT ACT SITE 465 465 BY SIMILARITY.  
CC FT ACT SITE 466 466 BY SIMILARITY.  
CC SEQUENCE 666 AA; 77095 MW; BA538EBD93DE9868 CRC64;  
  
Query Match 81.8%; Score 27; DB 1; Length 666;  
Best Local Similarity 85.7%; Pred. No. 86;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ORANLRA 7  
DB 486 QANLRA 492  
  
RESULT 13  
GLGB\_VIBU STANDARD; PRT; 715 AA.  
ID -GLGB\_VIBU  
AC Q8D4P0;

DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DE 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE 1,4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching  
 DE enzyme) (BE) (1,4-alpha-D-glucan:1,4-alpha-D-glucan 6-glucosyl-  
 DE transferase).  
 GN GLGB OR VV21252.  
 OS Vibrio vulnificus.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
 OC Vibrionaceae; Vibrio.  
 OX NCBI\_TaxID=672;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CMCP6;  
 RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,  
 RA Choy H.E.;  
 RT "Complete genome sequence of *Vibrio vulnificus* CMCP6.";  
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Catalyzes the formation of the alpha-1,6-glucosidic  
 CC linkages in glycogen by scission of a 1,4-alpha-linked  
 CC oligosaccharide from growing alpha-1,4-glucan chains and the  
 CC subsequent attachment of the oligosaccharide to the alpha-1,6  
 CC position (By similarity).  
 CC -!- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of  
 CC glycogen.  
 CC -!- PATHWAY: Glycogen biosynthesis; third step.  
 CC -!- SUBUNIT: Monomer (By similarity).  
 CC -!- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.  
 CC  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 DR EMBL; A5016812; AAO08146.1; -.  
 DR HAMAP; MF\_00685; -; 1.  
 DR InterPro; IPR006589; Alp\_amy1\_cat\_sub.  
 DR InterPro; IPR006047; Alpha\_amy1\_cat.  
 DR InterPro; IPR006407; GLGB.  
 DR InterPro; IPR004193; Glyco\_hydro\_13N.  
 DR Pfam; PF00128; alpha-amy1ase; 1.  
 DR Pfam; PF02922; isomylase\_N; 1.  
 DR SMART; SM00642; Amy; 1.  
 DR TIGRFAMs; TIGR01515; branching enzym; 1.  
 DR Glycogen biosynthesis; Transferase; Glycosyltransferase;  
 KW Complete proteome.  
 FT ACT\_SITE 291 291 BY SIMILARITY.  
 FT ACT\_SITE 326 326 BY SIMILARITY.  
 FT ACT\_SITE 331 331 BY SIMILARITY.  
 FT ACT\_SITE 331 331 BY SIMILARITY.  
 FT ACT\_SITE 394 394 BY SIMILARITY.  
 FT ACT\_SITE 396 396 BY SIMILARITY.  
 FT ACT\_SITE 449 449 BY SIMILARITY.  
 FT ACT\_SITE 516 516 BY SIMILARITY.  
 FT ACT\_SITE 517 517 BY SIMILARITY.  
 SQ SEQUENCE 715 AA; 81927 MW; 304C61BECF5A5F43 CRC64;  
 Query Match 81.8%; Score 27; DB 1; Length 715;  
 Best Local Similarity 85.7%; Pred. No. 93;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 QANLRA 7  
 Db 537 QANLRA 543  
 RESULT 14  
 ID GLGB VIBPA STANDARD; PRT; 755 AA.  
 AC Q87FRC;  
 DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE 1,4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching  
 DE enzyme) (BE) (1,4-alpha-D-glucan:1,4-alpha-D-glucan 6-glucosyl-  
 DE transferase).  
 GN GLGB OR VPA1618.  
 OS Vibrio parahaemolyticus.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
 OC Vibrionaceae; Vibrio.  
 OX NCBI\_TaxID=670;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RIMD 2210633 / Serotype O3:K6;  
 RA MEDLINE=22508454; PubMed=12620739;  
 RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,  
 RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,  
 RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;  
 RT "Genome sequence of *Vibrio parahaemolyticus*: a pathogenic mechanism  
 RT distinct from that of *V. cholerae*.";  
 RL Lancet 361:743-749(2003).  
 CC -!- FUNCTION: Catalyzes the formation of the alpha-1,6-glucosidic  
 CC linkages in glycogen by scission of a 1,4-alpha-linked  
 CC oligosaccharide from growing alpha-1,4-glucan chains and the  
 CC subsequent attachment of the oligosaccharide to the alpha-1,6  
 CC position (By similarity).  
 CC -!- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of  
 CC glycogen.  
 CC -!- PATHWAY: Glycogen biosynthesis; third step.  
 CC -!- SUBUNIT: Monomer (By similarity).  
 CC -!- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.  
 CC  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 DR EMBL; AP005089; BAC62961.1; -.  
 DR HAMAP; MF\_00685; -; 1.  
 DR InterPro; IPR006047; Alpha\_amy1\_cat.  
 DR InterPro; IPR004193; Glyco\_hydro\_13N.  
 DR Pfam; PF00128; alpha-amy1ase; 1.  
 DR Pfam; PF02922; isomylase\_N; 1.  
 DR Glycogen biosynthesis; Transferase; Glycosyltransferase;  
 KW Complete proteome.  
 FT ACT\_SITE 330 330 BY SIMILARITY.  
 FT ACT\_SITE 365 365 BY SIMILARITY.  
 FT ACT\_SITE 370 370 BY SIMILARITY.  
 FT ACT\_SITE 433 433 BY SIMILARITY.  
 FT ACT\_SITE 435 435 BY SIMILARITY.  
 FT ACT\_SITE 488 488 BY SIMILARITY.  
 FT ACT\_SITE 555 555 BY SIMILARITY.  
 FT ACT\_SITE 556 556 BY SIMILARITY.  
 SQ SEQUENCE 755 AA; 87504 MW; 8F052368B657D79 CRC64;  
 Query Match 81.8%; Score 27; DB 1; Length 755;  
 Best Local Similarity 85.7%; Pred. No. 98;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 QANLRA 7  
 Db 576 QANLRA 582  
 RESULT 15  
 ID KAB7 YEAST STANDARD; PRT; 1356 AA.  
 AC F31374;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Probable serine/threonine-protein kinase YAL017W (EC 2.7.1.1.-).  
GN YAL017W OR YAL002 OR FUN31.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288c / AB972;  
RX MEDLINE=93209532; PubMed=8458570;  
RA Ouellette B.F.F., Clark M.W., Keng T., Storms R.K., Zhong W.W.,  
RA Quelette B.F.F., Clark M.W., Barton A.B., Kaback D.B., Bussey H.;  
RA "Sequencing of chromosome I from Saccharomyces cerevisiae: analysis  
RT of a 32 kb region between the LTE1 and SPO7 genes.";  
RL Genome 36:32-42(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288c / AB972;  
RX MEDLINE=92221690; PubMed=1561836;  
RA Clark M.W., Zhong W.W., Keng T., Storms R.K., Barton A.,  
RA Kaback D.B., Bussey H.;  
RA "Identification of a Saccharomyces cerevisiae homolog of the SNF2  
RT transcriptional regulator in the DNA sequence of an 8.6 kb region in  
RT the LTE1-CYS1 interval on the left arm of chromosome I.";  
RL Yeast 8:133-145(1992).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288c / AB972;  
RX MEDLINE=93311122; PubMed=8322517;  
RA Clark M.W., Zhong W.W., Keng T., Storms R.K., Ouellette B.F.F.,  
RA Barton A., Kaback D.B., Bussey H.;  
RT "The YAL017 Gene on the left arm of chromosome I of Saccharomyces  
RT cerevisiae encodes a putative serine/threonine protein kinase.";  
RL Yeast 9:543-549(1993).  
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases..  
CC -!- SIMILARITY: Contains 2 PAS (PER-ARNT-SIM) dimerization domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; L05146; AAC04940.1; ..  
DR PIR; S33653; S33653.  
DR KSSP; Q63450; IA06.  
DR GerMOnline; I38359; ..  
DR SGD; S0000015; FUN31.  
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IMP.  
DR GO; GO:0019318; P:hexose metabolism; IGI.  
DR GO; GO:0006468; P:protein amino acid phosphorylation; IMP.  
DR InterPro; IPR000014; PAS domain.  
DR InterPro; IPR000719; Prot\_kinase.  
DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
DR Pfam; PF00069; pkinase; 1.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR SMART; SM00091; PAS; 2.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; FALSE\_NEG.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
KW Hypothetical protein; Transferase; Serine/threonine-protein kinase;  
KW ATP-binding; Repeat.  
FT DOMAIN 450 518 PAS 1.  
FT DOMAIN 738 807 PAS 2.  
FT DOMAIN 1096 1354 PROTEIN\_KINASE.  
FT NP\_BIND 1102 1110 ATP (BY SIMILARITY).  
FT BINDING 1125 1125 ATP (BY SIMILARITY).  
FT ACT\_SITE 1230 1230 BY SIMILARITY.  
SQ SEQUENCE 1356 AA; 152330 MW; 11B5F6F534B83986 CRC64;

Query Match 81.8%; Score 27; DB 1; Length 1356;

Best Local Similarity 71.4%; Pred. No. 1.9e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 QRANLEA 7  
Db 334 QRANLXS 340  
Search completed: June 9, 2004, 14:57:32  
Job time : 12 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2004, 14:54:35 ; Search time 40 Seconds  
(without alignments)  
55.216 Million cell updates/sec

Title: US-10-080-100-46  
Perfect score: 33  
Sequence: 1 QANLRA 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL.25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	100.0	599	4 Q9HCE8	Q9hce8 homo sapien
2	30	90.9	274	2 Q9S3U2	Q9s3u2 zymomonas m
3	30	90.9	556	16 Q889G6	Q889g6 pseudomonas
4	29	87.9	160	10 Q9SQH1	Q9sqh1 arachis hyp
5	29	87.9	245	16 Q8PQY7	Q8pqy7 xanthomonas
6	29	87.9	280	16 Q8D767	Q8d767 vibrio vuln
7	29	87.9	573	16 Q81CT8	Q81ct8 bacillus ce
8	29	87.9	748	16 Q89S99	Q89s99 bradyrhizob
9	29	87.9	804	16 Q7X656	Q7x656 oryza sativ
10	29	87.9	823	10 Q7XXV6	Q7xxv6 oryza sativ
11	29	87.9	877	10 Q7XXH6	Q7xxh6 oryza sativ
12	29	87.9	916	10 Q8L4K9	Q8l4k9 oryza sativ
13	29	87.9	1822	2 Q8VRG6	Q8vrg6 chloroflexu
14	28	84.8	97	9 Q857L9	Q857l9 mycobacteri
15	28	84.8	149	13 Q7ZZU3	Q7zzu3 oreochromis
16	28	84.8	175	4 Q9NZ85	Q9nz85 homo sapien

17	28	84.8	192	10 Q8LK83	Q8lk83 brassica ol
18	28	84.8	217	9 Q856W9	Q856w9 mycobacteri
19	28	84.8	399	16 Q8YDZ8	Q8ydz8 brucella me
20	28	84.8	439	4 Q8TBT4	Q8tbt4 homo sapien
21	28	84.8	440	4 Q96ES8	Q96es8 homo sapien
22	28	84.8	440	4 Q9GZM9	Q9gzm9 homo sapien
23	28	84.8	441	4 Q9BX56	Q9bx56 homo sapien
24	28	84.8	483	16 Q9RUB8	Q9rub8 deinococcus
25	28	84.8	496	13 Q42434	Q42434 notophthalm
26	28	84.8	498	13 Q9PV92	Q9pv92 brachydanio
27	28	84.8	507	13 Q7ZXU4	Q7zxu4 xenopus lae
28	28	84.8	511	13 Q90ZF7	Q90zf7 rana catesb
29	28	84.8	512	16 Q8FXK0	Q8fxk0 brucella su
30	28	84.8	515	13 Q8JFG4	Q8jfg4 oncorhynch
31	28	84.8	558	13 Q9PUB5	Q9pub5 brachydanio
32	28	84.8	595	11 Q8K030	Q8k030 mus musculu
33	28	84.8	673	15 Q89816	Q89816 mus dunni e
34	28	84.8	683	11 Q8B115	Q8b115 mus musculu
35	28	84.8	695	11 Q8C091	Q8c091 mus musculu
36	28	84.8	1581	4 Q9P2P5	Q9p2p5 homo sapien
37	28	84.8	3038	3 Q9Y8A5	Q9y8a5 aspergillus
38	27	81.8	93	16 Q87CU5	Q87cu5 xylella fas
39	27	81.8	98	9 Q9B0C0	Q9b0c0 mycobacteri
40	27	81.8	103	11 Q8END0	Q8end0 mus musculu
41	27	81.8	117	16 Q88N60	Q88n60 pseudomonas
42	27	81.8	148	16 Q87X08	Q87x08 pseudomonas
43	27	81.8	182	10 Q7XW65	Q7xw65 oryza sativ
44	27	81.8	188	16 Q7UR39	Q7ur39 rhodopirell
45	27	81.8	209	16 Q82EK7	Q82ek7 streptomyce

## ALIGNMENTS

## RESULT 1

Q9HCE8 PRELIMINARY; PRT; 599 AA.

AC Q9HCE8; 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Hypothetical protein KIAA1624 (Fragment).  
GN KIAA1624.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20450683; PubMed=10997877;  
RA Nagase T., Kikuno R., Nakayama M., Hirotsawa M., Ohara O.;  
RT XVIII. The complete sequences of 100 new cDNA clones from brain which  
RT code for large proteins in vitro."  
RL DNA Res. 7:273-281(2000).  
DR EMBL; AB046844; BAB13450.1; -  
DR Genbank; F017830; GPR107.  
KW Hypothetical protein.  
FT NON\_TER  
SQ SEQUENCE 599 AA; 66906 MW; DC2AD69FAE078500 CRC64;

Query Match 100.0%; Score 33; DB 4; Length 599;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QANLRA 7  
| | | | |  
Db 443 QANLRA 449

## RESULT 2

Q9S3U2 PRELIMINARY; PRT; 274 AA.

AC Q883U2;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Tryptophan synthase alpha subunit (EC 4.2.1.20) (Tryptophan synthase  
DE alpha chain).  
DE TRPA.  
OS Zymomonas mobilis (subsp. pomaceae).  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;  
OC Sphingomonadaceae; Zymomonas.  
OX NCBI\_TaxID=120044;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 29192;  
RA Eddy C.K., Ingram L.O.;  
RT "Zymomonas trpFBA genes";  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: THE ALPHA SUBUNIT IS RESPONSIBLE FOR THE ALDOL CLEAVAGE  
CC OF INDOLEGLYCEROL PHOSPHATE TO INDOLE AND GLYCERALDEHYDE 3-  
CC PHOSPHATE (BY SIMILARITY).  
CC -!- CATALYTIC ACTIVITY: L-SERINE + 1-(INDOL-3-YL)GLYCEROL 3-PHOSPHATE  
CC = L-TRYPTOPHAN + GLYCERALDEHYDE 3-PHOSPHATE + H(2)O.  
CC -!- PATHWAY: TRYPTOPHAN BIOSYNTHESIS; FIFTH (LAST) STEP.  
CC -!- SUBUNIT: Tetramer of two alpha and two beta chains (by  
CC similarity).  
CC -!- SIMILARITY: BELONGS TO THE TRPA FAMILY.  
DR EMBL; AF173835; AAD51339.1; -.  
DR HSSP; P00929; 2WSY.  
DR GO; GO:0016829; P:lyase activity; IEA.  
DR GO; GO:0004834; P:tryptophan synthase activity; IEA.  
DR GO; GO:0000162; P:tryptophan biosynthesis; IEA.  
DR InterPro; IPR003009; FMN enzyme.  
DR InterPro; IPR002028; Trp\_synthase.  
DR Pfam; PF00290; trp\_synthA; 1.  
DR ProDom; PD001535; Trp\_synthaseA; 1.  
DR TIGRFAMs; TIGR00262; trpA; 1.  
DR PROSITE; PS00167; TRP SYNTHASE ALPHA; 1.  
KW Lyase: Tryptophan biosynthesis.  
SQ SEQUENCE 274 AA; 28840 MW; A1971BB93D29A324 CRC64;  
  
Query Match 90.9%; Score 30; DB 2; Length 274;  
Best Local Similarity 85.7%; Pred. No. 51;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 QXANLRA 7  
Db 65 QXANLRA 71  
  
RESULT 3  
Q889G6 PRELIMINARY; PRT; 656 AA.  
AC Q889G6;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE TonB-dependent siderophore receptor, putative.  
GN PSPT00784.  
OS Pseudomonas syringae (pv. tomato).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=323;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DC3000;  
RA Buell R., Joardar V., Khouri H., Fedorova N., Tran B., Russell D.,  
RA Berry K., Uterback T., Van Aken S., Feldblyum T., Gwin M.,  
RA Dodson R., DeBoy R., Durkin A., Kolonay J., Madupu R., Daugherty S.,  
RA Brinkac L., Beanan M., Haft D., Selengut J., Nelson W., Davidson T.,  
RA White O., Fraser C., Collier A.;  
RT "Complete sequence of Pseudomonas syringae";  
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE016858; AAO54326.1; -.

DR TIGR; PSPT00784; -;  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR GO; GO:0005215; F:transporter activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR000531; TonB\_boxC.  
DR Pfam; PF00593; TonB\_dep\_Rec; 1.  
DR PROSITE; PS01156; TONB\_DEPENDENT\_REC\_2; 1.  
KW Receptor; Complete proteome.  
SQ SEQUENCE 656 AA; 72591 MW; 2238CC4E57A63008 CRC64;  
  
Query Match 90.9%; Score 30; DB 16; Length 656;  
Best Local Similarity 85.7%; Pred. No. 1.3e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 QXANLRA 7  
Db 566 QXANLRA 572  
  
RESULT 4  
Q89QH1 PRELIMINARY; PRT; 160 AA.  
AC Q89QH1;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Allergen.  
DE ARA H 7.  
OS Arachis hypogaea (Peanut).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;  
OC Arachis.  
OX NCBI\_TaxID=3818;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Virginia; TISSUE=Seed;  
RX MEDLINE=99406463; PubMed=10474031;  
RA Kleber-Janke T., Cramer R., Appenzeller U., Schlaak M., Becker W.M.;  
RT "Selective cloning of peanut allergens, including profilin and 2S  
RT albumins, by phage display technology";  
RL Int. Arch. Allergy Immunol. 119:265-274(1999).  
DR EMBL; AF091737; AAD56719.1; -.  
DR InterPro; IPR003612; AAI.  
DR Pfam; PF00234; tryp\_alpha\_amyl; 1.  
DR SMART; SM00499; AAI; 1.  
SQ SEQUENCE 160 AA; 18417 MW; 9F9E4CEE68808D4C CRC64;  
  
Query Match 87.9%; Score 29; DB 10; Length 160;  
Best Local Similarity 100.0%; Pred. No. 50;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 QXANLRA 6  
Db 50 QXANLRA 55  
  
RESULT 5  
Q89QV7 PRELIMINARY; PRT; 245 AA.  
AC Q89QV7;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE ABC transporter ATP-binding protein.  
GN YEHX OR XAC0182.  
OS Xanthomonas axonopodis (pv. citri).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
OC Xanthomonadaceae; Xanthomonas.  
OX NCBI\_TaxID=92829;  
RN [1]  
RP SEQUENCE FROM N.A.

RC	STRAIN=306 / ATCC 13902 / XV 101;
RX	MEDLINE=22022145; PubMed=12024217;
RA	da Silva A.C.R., Ferro J.A.O., Reinach F.C., Farah C.S., Furian L.R.,
RA	Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA	Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA	Camarote G., Cannavon F., Cardozo J., Chambergo F., Clapina L.P.,
RA	Ciccarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA	Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA	Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA	Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA	Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA	Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA	Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA	Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA	Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA	Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA	Setubal J.C., Kitajima J.P.;
RT	"Comparison of the genomes of two Xanthomonas pathogens with differing
RT	host specificities.";
RL	Nature 417:459-463(2002).
RL	EMBL: AF011642; AA035074.1; -.
DR	GO: GO:0016020; C:membrane; IEA.
DR	GO: GO:0005224; F:ATP binding; IEA.
DR	GO: GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . ; IEA.
DR	GO: GO:0006810; P:transport; IEA.
DR	InterPro: IPR003439; ABC transporter.
DR	Pfam: PF00005; ABC tran; 1.
DR	PROSITE: PS00211; ABC_TRANSPORTER_1; 1.
DR	PROSITE: PS00893; ABC_TRANSPORTER_2; 1.
KW	ATP-binding; Complete proteome.
SQ	SEQUENCE 245 AA; 26698 MW; B9BC33B69C48D52D CRC64;
Query Match 87.9%; Score 29; DB 16; Length 245;	
Best Local Similarity 85.7%; Pred.No. 79;	
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;	
QY	1 QRANLRA 7      :
Db	66 QRANLQA 72
RESULT 6	
Q8D767	PRELIMINARY; PRT; 280 AA.
ID	Q8D767
AC	Q8D767;
DT	01-MAR-2003 (TrEMBLrel. 23, Created)
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE	Outer membrane receptor protein.
GN	VV20298.
OS	Vibrio vulnificus.
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC	Vibrionaceae; Vibrio.
OX	NCBI_TaxID=672;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=GMCPE;
RA	Rhee J.H., Kim S.Y., Chung S.-S., Kim J.J., Moon Y.H., Jeong H.,
RA	Choy H.E.;
RT	"Complete genome sequence of <i>Vibrio vulnificus</i> CMCP6.";
RT	Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR	EMBL: AE016809; AAC07260.1; -.
DR	GO: GO:0004872; F:receptor activity; IEA.
KW	Receptor; Complete proteome.
SQ	SEQUENCE 280 AA; 31390 MW; 1E2C85898FE1C7A4 CRC64;
Query Match 87.9%; Score 29; DB 16; Length 280;	
Best Local Similarity 100.0%; Pred.No. 91;	
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 QRANLR 6 
Db	25 QRANLR 30

```

RL DNA Res. 9:189-197(2002).
DR EMBL; AF005944; BAC47771.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0006935; P:chemotaxis; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR004089; Charaxis_trans.
DR InterPro; IPR003650; HAMP.
DR InterPro; IPR000727; T_SNARE.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF00015; MCPsignal; 1.
DR PROSITE; PS01111; CHEMOTAXIS_TRANSDUC_2; 1..
DR PROSITE; PS00885; HAMP; 1.
DR PROSITE; PS01192; T_SNARE; 1.
KW Complete proteome.
SQ SEQUENCE 748 AA; 79622 MW; 718ECFA61FDAC892 CRC64;

Query Match 87.9%; Score 29; DB 16; Length 748;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRANLR 6
DB 144 QRANLR 149

RESULT 9
ID Q7X656 PRELIMINARY; PRT; 804 AA.
AC Q7X656;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE OSJNB0067G11.2 protein (OSJNB0115109.24 protein).
GN OSJNB0067G11.2 OR OSJNB0115109.24.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,
RA Liu Y.L., Mu J., Yu Z., Chen L., Fan D.L., Weng Q.J., Zhang L.,
RA Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C., Li T.,
RA Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen Z.H.,
RA Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,
RA Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,
RA Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,
RA Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,
RA Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.;
RA Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AL731586; CAD40389.1; -.
SQ SEQUENCE 823 AA; 89530 MW; 27A27E890AA48787 CRC64;

Query Match 87.9%; Score 29; DB 10; Length 823;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRANLR 6
DB 249 QRANLR 254

RESULT 11
ID Q7XKH6 PRELIMINARY; PRT; 877 AA.
AC Q7XKH6;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE OSJNB0033P05.11 protein.
GN OSJNB0033P05.11.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,
RA Liu Y.L., Mu J., Yu Z., Chen L., Fan D.L., Weng Q.J., Zhang L.,
RA Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C., Li T.,
RA Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen Z.H.,
RA Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,
RA Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,
RA Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,
RA Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,
RA Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.;
RA Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AL731634; CAE05672.1; -.
SQ SEQUENCE 877 AA; 97395 MW; BCSAA7F6B8593158 CRC64;

Query Match 87.9%; Score 29; DB 10; Length 877;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;

QY 1 QRANLR 6
DB 249 QRANLR 254

RESULT 11
ID Q7XKH6 PRELIMINARY; PRT; 877 AA.
AC Q7XKH6;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE OSJNB0033P05.11 protein.
GN OSJNB0033P05.11.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,
RA Liu Y.L., Mu J., Yu Z., Chen L., Fan D.L., Weng Q.J., Zhang L.,
RA Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C., Li T.,
RA Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen Z.H.,
RA Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,
RA Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,
RA Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,
RA Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,
RA Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.;
RA Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AL731634; CAE05672.1; -.
SQ SEQUENCE 877 AA; 97395 MW; BCSAA7F6B8593158 CRC64;

Query Match 87.9%; Score 29; DB 10; Length 877;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;

QY 1 QRANLR 6
DB 249 QRANLR 254

```

	Matches	6;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	QRANLR	6							
Db	240	QRANLR	245							
RESULT	12									
Q8L4K9										
ID	Q8L4K9	PRELIMINARY;	PRT;	916	AA.					
AC	Q8L4K9;									
DT	01-OCT-2002	(TrEMBLrel. 22, Created)								
DT	01-OCT-2002	(TrEMBLrel. 22, Last sequence update)								
DT	01-OCT-2003	(TrEMBLrel. 25, Last annotation update)								
DE	Hypothetical protein.									
GN	OSUNBA0079H13.17	OR OSJNEB0038H12.10.								
OS	Oryza sativa	(japonica cultivar-group).								
OC	Eukaryota:	Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;								
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;									
OC	Erbartoideae; Oryzeae; Oryza.									
OX	NCBI_TaxId=39947;									
RN	[1]									
RP	SEQUENCE FROM N.A.									
RC	STRAIN=cv. Nipponbare;									
RA	Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Kim M.M.,									
RA	Overton II L.L., Tsitrin T., Bera J.J., Tsitrin T., Krol M.I., Jarrahi B.B.,									
RA	Jin S.S., Koo H., Zismann V., Hsiao J., Blunt S., Vanaken S.S.,									
RA	Utterback T.T., Feldblum T.V., Yang Q.Q., Haas B.J., Suh B.B.,									
RA	Peterson J.J., Quackenbush J., White O., Salzberg S.L., Fraser C.M.;									
RL	"Oryza sativa chromosome 10 BAC OSUNBA0079H13 genomic sequence.";									
RL	Submitted (SEP-2002) to the EMBL/GenBank/DBSJ databases.									
RN	[2]									
RP	SEQUENCE FROM N.A.									
RC	STRAIN=cv. Nipponbare;									
RA	Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,									
RA	Overton II L.L., Tsitrin T., Kim M.M., Bera J.J., Jin S.S.,									
RA	Fadrosch D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,									
RA	Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblum T.V.,									
RA	Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,									
RA	White O., Salzberg S.L., Fraser C.M.;									
RL	"Oryza sativa chromosome 10 BAC OSJNEB0038H12 genomic sequence.";									
RL	Submitted (SEP-2002) to the EMBL/GenBank/DBSJ databases.									
RN	[3]									
RP	SEQUENCE FROM N.A.									
RC	STRAIN=cv. Nipponbare;									
RA	The Rice Chromosome 10 Sequencing Consortium;									
RT	"In-depth view of structure, activity, and evolution of rice									
RT	chromosome 10,";									
RL	Science 300:1568-1569(2003).									
RN	[4]									
RP	SEQUENCE FROM N.A.									
RC	STRAIN=cv. Nipponbare;									
RA	Buell C.R., Wang R.A., McCombie W.R., Messing J., Yuan Q.;									
RL	Submitted (MAY-2003) to the EMBL/GenBank/DBSJ databases.									
DR	EWEL; AC104616; AAM54158.1;-									
DR	EWEL; AC105932; AAM04984.1;-									
DR	EWEL; AE017067; AAP52540.1;-									
DR	Gramene; Q8L4K9; -									
DR	InterPro; IPR002559; Transposase_11.									
DR	Pfam; PF01609; Transposase_11; 1-									
KW	Hypothetical protein.									
SQ	SEQUENCE	916	AA;	101323	MW;	6D72ECCEE981433B	CRC64;			
Query Match	87.9%; Score 29; DB 10; Length 916;									
Best Local Similarity	100.0%; Pred.No.3.2e+02;									
Matches	6;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;	
Qy	1	QRANLR	6							
Db	249	QRANLR	254							
RESULT	13									

DR GO; GO:0004519; F:endonuclease activity; IEA.  
 DR GO; GO:0003676; F:nucleic acid binding; IEA.  
 DR InterPro; IPR002711; HNH.  
 DR Pfam; PF01844; HNH; 1.  
 SQ SEQUENCE 97 AA; 11190 MW; D3FA82B91C2B2F3E CRC64;

Query Match 84.8%; Score 28; DB 9; Length 97;  
 Best Local Similarity 100.0%; Pred. No. 52;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QANLEA 7  
 |||||  
 DB 56 QANLEA 61

RESULT 15

Q7ZZU3 PRELIMINARY; PRT; 149 AA.  
 AC Q7ZZU3;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Keratin 8 (Fragment).  
 OS Oreochromis mossambicus (Mozambique tilapia) (Tilapia mossambica).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidel;  
 OC Cichlidae; Oreochromis.  
 OX NCBI\_TaxID=8127;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lo M.J., Chan L.W., Weng C.F.;  
 RT "The functional genomic assay of tilapia larvae during development.";  
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY260748; RAP22041.1; ..  
 DR GO; GO:0005882; C:intermediate filament; IEA.  
 DR InterPro; IPR001664; IF.  
 DR InterPro; IPR003054; Keratin\_II.  
 DR Pfam; PF00038; Filament; 1.  
 DR PRINTS; PR01276; TYPE2KERATIN.  
 DR PROSITE; PS00226; IF; 1.  
 KW Keratin.  
 FT NON\_TER 1  
 FT NON\_TER 149  
 SQ SEQUENCE 149 AA; 16502 MW; FDC79BB1016C14D9 CRC64;

Query Match 84.8%; Score 28; DB 13; Length 149;  
 Best Local Similarity 85.7%; Pred. No. 82;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QANLEA 7  
 |||||  
 DB 45 QANLEA 51

Search completed: June 9, 2004, 14:58:25  
 Job time: 41 secs